

Rigorous RG algorithms and area laws for low energy eigenstates in 1D

Itai Arad*

Zeph Landau†

Umesh Vazirani‡

Thomas Vidick§

Abstract

One of the central challenges in the study of quantum many-body systems is the complexity of simulating them on a classical computer. A recent advance [LVV15] gave a polynomial time algorithm to actually compute a succinct classical description for unique ground states of gapped 1D quantum systems. Despite this progress many questions remained unresolved, including whether there exist rigorous efficient algorithms when the ground space is degenerate (and $\text{poly}(n)$ dimensional), or for the $\text{poly}(n)$ lowest energy states for 1D systems, or even whether such states admit succinct classical descriptions or area laws.

In this paper we give a new algorithm for finding low energy states for 1D systems, based on a rigorously justified RG type transformation. In the process we resolve some of the aforementioned open questions, including giving a polynomial time algorithm for $\text{poly}(n)$ degenerate ground spaces and an $n^{O(\log n)}$ algorithm for the $\text{poly}(n)$ lowest energy states for 1D systems (under a mild density condition). We note that for these classes of systems the existence of a succinct classical description and area laws were not rigorously proved before this work. The algorithms are natural and efficient, and for the case of finding unique ground states for frustration-free Hamiltonians the running time is $\tilde{O}(nM(n))$, where $M(n)$ is the time required to multiply two $n \times n$ matrices.

1 Introduction

One of the central challenges in the study of quantum systems is their exponential complexity [Fey82]: the state of a system on n particles is given by a vector in an exponentially large Hilbert space, so even giving a classical description (of size polynomial in n) of the state is a formidable challenge. A priori, the task is not impossible, as the physically relevant states lie in a tiny corner of the Hilbert space. To be useful, the classical description of these states must support the efficient computation of expectation values of local observables. The renormalization group formalism [Wil75] provides a sweeping approach to this corner of Hilbert Space by suggesting that physically relevant quantum states can be coarse-grained at different levels of granularity, or length scales, thereby iteratively eliminating the irrelevant degrees of freedom. Ideally, doubling the length scale in such a coarse-graining process does not increase the size of the description, as it only retains physically relevant degrees of freedom. This lies at the core of Wilson’s numerical renormalization group (NRG) approach that successfully solved the Kondo problem [Wil75]. The approach was subsequently improved by White [Whi92, Whi93], to obtain the famous Density Matrix Renormalization Group (DMRG)

*Centre for Quantum Technologies (CQT), National University of Singapore, Singapore

†Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, U.S.A.

‡Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, U.S.A.

§Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, USA. email: vidick@cms.caltech.edu

algorithm [Whi92, Whi93], which is widely used for solving for the ground and low energy states of 1D systems.

Formally understanding the success of DMRG (and NRG) has been extremely challenging, since it touches on deep questions about how non-local correlations such as entanglement arise from local interaction Hamiltonians. This tension between the locality of interactions and non-local correlations is manifested in quantum phase transitions, and gives rise to a rich set of quantum phases of matter. While the primary goal of this paper is to present rigorous new results about the nature of entanglement in low-energy states of 1D systems, along with efficient classical algorithms for solving such systems, the techniques introduced here also shed new light on the Renormalization Group (RG) framework.

Central to our approach is a rigorously justified RG transformation that relies on a carefully constructed subclass of Approximate Ground Space Projections (AGSPs). In general, given a ground space T , AGSPs are linear operators that preserve T (or a nearby space) while shrinking the orthogonal complement to T . The AGSPs used for the RG transformation we describe here require a very stringent tradeoff between their entanglement rank and the amount of shrinking they achieve (essentially the condition $D\Delta < 1$ in the AGSPs of Refs. [ALV12, AKLV13]). The RG transformation instantiates a coarse graining approach that allows us to ‘merge’ adjacent blocks of particles, creating larger and larger blocks without increasing the effective degrees of freedom: at each length scale there is a representation of the low-energy spectrum of H involving only a small number of states of moderate bond dimension. Specifically, suppose a given 1D system has a low-energy subspace T of dimension r . And suppose that for any block of ℓ particles, there is a subspace S of dimension $r^{1+o(1)}$ of the associated Hilbert space such that the reduced density matrix of every state in T on the block is well-approximated by states from S . Then we provide a prescription, based on the use of our AGSP operator, for merging two such neighboring subspaces into a single subspace S' for a block of 2ℓ particles, such that S' has the *same dimension* $r^{1+o(1)}$ and satisfies the *same approximation quality* for the subspace T . This is similar in spirit to what the DMRG algorithm is aiming for, except that here the merging process is rigorously controlled and as a result there is no need for the iterative and heuristic left-and-right sweeping procedure that is used in DMRG to improve the overlap with the ground state.

This new RG formalism directly leads to succinct MPS descriptions in the following two settings, previously completely open:

1. *Hamiltonians with a Degenerate Gapped ground space (DG)*: H has smallest eigenvalue ε_0 with associated eigenspace of dimension $r = \text{poly}(n)$, and second smallest eigenvalue ε_1 such that $\varepsilon_1 - \varepsilon_0 \geq \gamma$ with γ a constant termed the spectral gap. See Corollary 3.10 for a formal statement of the ensuing area law.
2. *Gapless Hamiltonians with a Low Density of low-energy states (LD)*: The dimension of the space of all eigenvectors of H with eigenvalue in the range $[\varepsilon_0, \varepsilon_0 + \eta]$, for some constant $\eta > 0$, is $r = \text{poly}(n)$. See Corollary 5.4 for a formal statement.

Further work results in efficient algorithms for finding the relevant states in both settings above, running in polynomial time for the first case and quasi-polynomial for the second. (See Theorem 4.2 for a precise statement in the first case, and Theorem 5.6 for the second case.) To a large extent, these algorithms can be viewed as *provably efficient* numerical RG algorithms.

These results should be understood in the context of a substantial body of prior work studying ground state entanglement in 1D systems. Central to this work is the so-called *area law for entanglement entropy* — a conjecture which states that in ground states of gapped local Hamiltonians the entanglement entropy of a region should scale as its surface area, rather than its volume (see, for example, the review article

Ref. [ECP10]). A landmark result by Hastings [Has07] proved this conjecture for gapped 1D systems with unique ground state. He also proved that for such systems the ground state can be well-approximated by a matrix product state (MPS) with bond dimension that scales polynomially in the number of spins in the system. This rigorously justified DMRG’s restriction of its search space to the manifold of succinct MPSs [ÖR95, RÖ97].

Hastings’ bound for the entanglement entropy scales as $e^{O(J\gamma^{-1} \log d)}$ for a chain of d -dimensional spins with nearest-neighbor interactions of strength J and spectral gap $\gamma > 0$. This bound is exponentially larger than the conjectured bound $O(J\gamma^{-1} \log d)$. A very different approach for proving the area law in 1D gapped systems with a unique ground state was pioneered in Refs. [ALV12, AKLV13], which gave a bound of $O(J\gamma^{-1} \log^3(dJ\gamma^{-1}))$ on the entanglement entropy. Central to this new proof was an AGSP with similar entanglement rank to shrinking tradeoff ($D\Delta < 1$) to that of this work. In addition to the improved bound this new proof also showed that the ground state can be well-approximated by an MPS with a bond dimension that scales as a *sublinear* function of the number of spins in the system, thereby giving a sub-exponential time algorithm for finding such a state. Subsequently, [LVV15] provided an efficient (polynomial time) algorithm that computes MPS descriptions for unique ground states of 1D gapped Hamiltonians. This provided a rigorous counterpoint to DMRG, formally establishing that the computational problem addressed by this heuristic is computationally tractable.

Many systems, even in one dimension, do not strictly satisfy the two essential conditions required for the application of these previous works: (i) the existence of a unique ground state, and (ii) a constant spectral gap above the ground state energy. For example, it may be that the system has a spectral gap, but the ground space is degenerate with polynomially bounded degeneracy — for instance, Ref. [dBOE10] consider a wide class of “natural” frustration-free local Hamiltonians in 1D for which the dimension of the ground space scales linearly with the number of particles. It is also interesting to consider the case of systems which display a vanishing gap (as the number of particles increases), while still maintaining a polynomial density of low-energy eigenstates (see for instance Ref. [KLW15]). The assumption of polynomial density arises naturally as one considers local perturbations of gapped Hamiltonians: while conditions under which the existence of a spectral gap remains stable are known [BHM10], it is expected that as the perturbation reaches a certain constant critical strength the gap will slowly close; in this scenario it is reasonable to expect that low-lying eigenstates should remain amenable to analysis.

Even though in all these scenario one would expect an area law, or a logarithmically corrected law, to hold, aside for specific cases where a direct computation is possible few rigorous results are known [ECP10]. Chubb and Flammia [CF15] extended the approach from Ref. [LVV15] to establish an area law (and efficient algorithm) for gapped Hamiltonians with a constant degeneracy in the ground space. Masanes [Mas09] proves an area law with logarithmic correction under a strong assumption on the density of states (slightly stronger than our (LD) assumption), together with an additional assumption on the exponential decay of correlations in the ground state.

The results presented in this paper address these gaps by proving area laws, succinct descriptions and efficient classical algorithms for low energy states in these settings. Moreover, the new algorithms are natural in the sense that they do not rely on the kind of ϵ -net discretization used in the algorithms from Refs. [LVV15, Hua14, CF15]. For the case of frustration-free Hamiltonians this results in a tightly-controlled running time: for systems with a unique ground state the algorithm runs in time $O(n^{1+o(1)}M(n))$, where $M(n) = O(n^{2.38})$ is the time required to multiply two $n \times n$ matrices.

Background and Methods Before describing our results we briefly review the state of the art. The 1D algorithm from Ref. [LVV15] starts with a partial solution on a small number of particles (d -dimensional

qudits) and proceeds iteratively, sweeping along the 1D chain from left to right. At each iteration the procedure extends the partial solution to include an additional particle. The challenge is to perform each iteration in polynomial time, as well as to keep the complexity of the description of the partial solution polynomially bounded. The 1D algorithm uses succinct representations of quantum states known as matrix product states (MPS). Roughly speaking, an MPS over n qudits of dimension d is specified by a set of d matrices of size $D \times D$ at each site, where the parameter D is called the bond dimension and determines the amount of entanglement the MPS can support. The iterative step of the algorithm of Ref. [AKLV13] relies on an ϵ -net to discretize a constant-radius ball in the linear space of these matrices at each site, and uses a dynamical-programming based approach inspired by [AAI10, SC10] to iteratively find the optimal set of matrices by performing an exhaustive search through the relevant net at each iteration. The size of the ϵ -net places a thorny limitation on this approach: the current best estimates for the size of D needed in order to obtain a constant approximation to the ground state scale as $D = d^{O(\log^{2/3} n)}$; since the net has dimension exponential in D its size scales super-polynomially. The solution given in Ref. [LVV15] is to use a constant D and prevent the error from accumulating by approximately projecting intermediate solutions on the ground state using an efficiently implementable but suitably relaxed version of an AGSP.¹

Since its publication the latter algorithm has been extended and improved in a number of ways. In Ref. [Hua14] the AGSP is replaced by a different construction based on ideas of Hastings [Has04, Has10], resulting in an improved dependence of the running time on the ground state energy. In Ref. [CF15] the algorithm is augmented to find a basis for a degenerate ground space with constant dimension. In both cases the basic structure of the algorithm, including the use of an ϵ -net to discretize the space of matrices, remains unchanged. Besides being somewhat unappealing aesthetically, this brute-force discretization through the ϵ -net is also a significant obstacle to adapting the algorithm into a practical heuristic. Even granting a constant D , its size scales exponentially with γ^{-1} , where γ is the spectral gap, which results in a doubly exponential in γ^{-1} scaling of the running time of the algorithm. Finally, all algorithms assume the existence of an underlying succinct description of the ground state(s) in the form of MPS (i.e. a form of an area law), limiting their applicability to the case where a sufficient area law is known to hold, i.e. 1D gapped Hamiltonians with a constant degeneracy in the ground space.

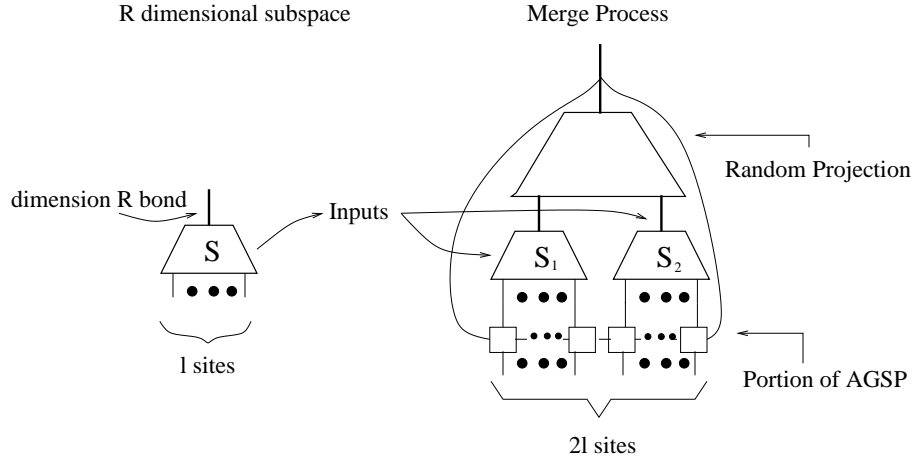
Although the RG-like approach introduced in this paper can be carried out by adding a particle in each iteration as above, it also naturally lends itself to working on all particles simultaneously in each iteration, combining adjacent blocks of particles to double their size. This allows us to proceed along a tree-like structure, reducing the number of iterations from n to $\log n$. As we shall see this has important implications for the use of this approach to prove area laws.

As described earlier, the basic step in our RG-like approach is a procedure, the *Merge process*, which takes as input two subspaces S_1 and S_2 of the Hilbert spaces associated to two adjacent blocks of ℓ particles such that $\dim(S_1), \dim(S_2) = \dim(T)^{1+o(1)}$, where T is the low-energy subspace of H . (For clarity we suppress dependence on parameters such as the local dimension d or the spectral gap γ ; see Proposition 3.1 for a more precise statement.) The procedure outputs a subspace S of dimension $\dim(T)^{1+o(1)}$ of the Hilbert space associated with the composite block of 2ℓ particles, such that if S_1 and S_2 approximately preserve the low energy subspace T , then so does S . Merge starts with the trivial subspace obtained as the tensor product $S_1 \otimes S_2$, whose dimension $\dim(S_1) \dim(S_2)$ is too large, and whose degree of approximation (suitably measured) is worse than desired by a factor of 2 (which is unsustainable in the long run). The first step of Merge consists in selecting a small random subspace of $S_1 \otimes S_2$. This has the effect of drastically reducing the dimension, to even below what is ultimately needed, at the expense of further blowing up the

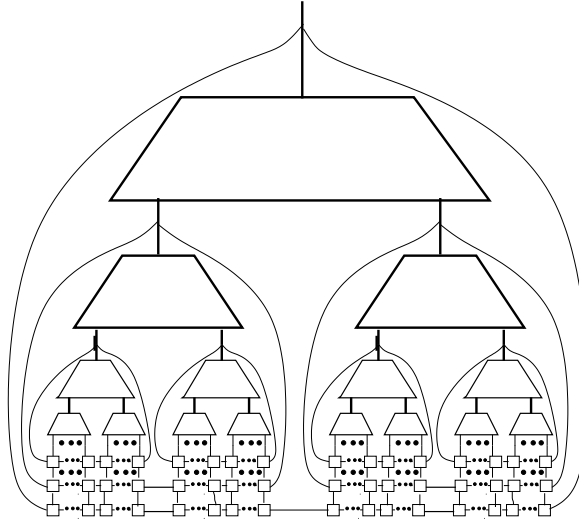
¹Unlike the core AGSP used in the RG framework here, this AGSP did not require a tight tradeoff between entanglement rank and shrinking.

approximation error. The second step controls the induced error by the application of a carefully chosen AGSP, improving the error while only mildly blowing up the dimension. The two-step process may be visualized as a (two stroke) pump: the first step (random projection) reduces the dimension, but leaves the dimension-error trade-off essentially unchanged. The second step (AGSP) improves the dimension-error tradeoff sufficiently so that the two steps together restore both dimension and error to their values at the start of the iteration.

A tensor network picture of the Merge process is provided in the figure below. Beginning with inputs representing subspaces of ℓ particles shown on the left, the Merge process (shown on the right) outputs a representation of a small subspace on 2ℓ particles.



The algorithm iteratively applies the Merge process, with the resulting tensor network picture after 4 steps given below.²



Our construction results in a partial isometry that is reminiscent of a MERA [Vid08, Vid09], and it should be fruitful to compare and contrast them as well as to standard RG. In particular, both our construction and

²We are grateful to Christopher T. Chubb for originally suggesting these pictures to us.

RG build subspaces in a binary tree fashion. However, whereas RG can be realized as a tensor network on a binary tree (where each node represents the partial isometry associated with selecting only a small portion of the previous space), the use of the AGSP in our construction allows for selection of the small subspace outside the tensor product of the previous two spaces.

We now indicate how an area law can be proven using this RG-like approach based on the Merge process, and then explain how it can be made algorithmically efficient. To prove an area law it (essentially) suffices to use the existing constructions of AGSPs from [AKLV13]. The main challenge in designing, and analyzing, the Merge process lies in selecting the right measure of approximation error for a subspace S . We accomplish this by expanding the notion of viable set to the setting of a target subspace T from the setting of a single target vector introduced in Ref. [LVV15]. The properties of such generalized viable sets play a central role in the analysis of our RG-like approach. We note that for the reader solely interested in the proof of an area law for the case of 1D gapped degenerate Hamiltonians Section 7 provides a self-contained proof that is stripped down to its bare essentials: the generalization of the definition of viable set along with the aforementioned “two stroke pump” procedure.

Designing an algorithm to find the kind of efficient representations of low energy states promised to exist by the newly proved area laws is more challenging. Besides the dimension of the subspace S , which corresponds to the number of quantum states stored by the algorithm, the running time is also affected by the description complexity (bond dimension) of states considered at intermediate steps of the algorithm. Controlling this bond dimension requires that the AGSP operators at the heart of the Merge process not only provide a favorable tradeoff between dimension and error, but can also be efficiently constructed and applied. A substantial effort in this paper is devoted towards the construction of a suitable class of AGSPs — termed spectral AGSPs — that meet all these conditions. This relies on modifying the *hard truncation* procedure used to control the norm of the local Hamiltonian in [AKLV13] to a procedure of *soft truncation*, which approximates the thresholding function with $x \mapsto t(1 - e^{-x/t})$. We crucially show that the latter function of a local Hamiltonian can be efficiently encoded as an MPO by leveraging recent results from [MSVC15] based on the use of the cluster expansion [Has06, KGK⁺14].

The polynomial factor increase in bond dimension over each iteration yields a quasi-polynomial bond dimension at the end of the $\log n$ iterations of the algorithm. This increase can be further controlled by including a step of bond trimming after each application of the AGSP used in the Merge process. The result is a polynomial-time algorithm for computing a basis for a $\text{poly}(n)$ gapped degenerate ground space. For the gapless case, a little more work is needed; our main idea consists in creating a slowly decreasing artificial spectral gap, which again lets us implement a suitable AGSP and yields a quasi-polynomial time algorithm.

As noted earlier our new algorithms could potentially be made very efficient. The main bottlenecks are the complexity of the AGSP and the MPS bond dimension that must be maintained. In the case of a frustration-free Hamiltonian with unique ground state we obtain a running time of $O(2^{O(1/\gamma^2)} n^{1+o(1)} M(n))$, where $M(n)$ is matrix multiplication time. This has an exponentially better scaling in terms of the spectral gap γ (due to avoidance of the ε -net argument) and saves a factor of $n/\log n$ (due to the logarithmic, instead of linear, number of iterations) as compared to an algorithm for the same problem considered in [Hua14]. We speculate that it might further be possible to limit the bond dimension of all MPS considered to $n^{o(1)}$ (instead of $n^{1+o(1)}$ currently), which, if true, would imply a nearly-linear time $O(n^{1+o(1)})$ algorithm.

The ideas in this paper were first reported at the Reunion Workshop for the Quantum Hamiltonian Complexity program at the Simons Institute [ref15]. In subsequent work [Hua15] Huang gave a simple algorithm restricted to the frustration-free with unique ground state case, yielding a different dependence on γ and n .

Organization of the paper. We start with some preliminaries and notation in Section 2. Section 3 introduces our main tool, the Merge process, and employs it to derive an area law for 1D gapped Hamiltonian with polynomial degeneracy in the ground space. In Section 4 we build on the approach to develop an efficient algorithm for the same systems. Section 5 extends these results to the case of gapless Hamiltonians satisfying a low-density of low-energy eigenstates assumption. In Section 6 we describe in detail the constructions of AGSP that underlie our results. Section 7 is a stand alone section that describes a short approach to proving our area law for the gapped degenerate case based on a bootstrapping argument. Appendices A, B and C contain proofs that were omitted from the main text.

2 Preliminaries and Notation

We begin by describing the basic setup for all our results.

Definition 2.1. Let $H = \sum_{i=1}^n h_i$ be a local Hamiltonian acting on the Hilbert space

$$\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d \otimes \dots \otimes \mathbb{C}^d \simeq (\mathbb{C}^d)^{\otimes n}$$

associated with a 1D chain of n qudits, each of local dimension d . Each h_i is assumed to be a non-negative operator with norm at most 1 acting on the i -th and $(i + 1)$ -st qudits. We denote by $\varepsilon_0 \geq 0$ the smallest eigenvalue (ground energy) of H , and consider the following three assumptions:

- (FF) **Frustration-Free:** H is frustration-free ($\varepsilon_0 = 0$) with a unique ground state $|\Gamma\rangle$ and a spectral gap $\gamma > 0$ above the ground state. In this case we let $T = \text{Span}\{|\Gamma\rangle\}$ denote the ground space of H .
- (DG) **Degenerate Gapped:** H has a degenerate ground space T of dimension $r = \text{poly}(n)$, along with a spectral gap $\gamma > 0$ above the ground space.
- (LD) **Low Density:** There is a constant $\eta > 0$ such that the space T spanned by eigenvectors of H with eigenvalue in the range $[\varepsilon_0, \varepsilon_0 + \eta]$ has dimension r at most $\text{poly}(n)$.

For $A \subseteq \{1, \dots, n\}$ we denote the Hilbert associated with the qudits in A by \mathcal{H}_A , e.g. $\mathcal{H}_{[1,3]} = \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ corresponds to the first three qudits. Separately, for any operator (Hamiltonian) H , $H_{[a,b]}$ will denote the subspace spanned by the eigenvectors of H with eigenvalues in the interval $[a, b]$. For a set S of vectors we denote by P_S the orthogonal projection onto the span of S and refer to $\dim(\text{Span}(S))$ as the size of S , denoted $|S|$. We often identify sets of vectors with the vector space they span.

We use standard $O(\cdot)$, $o(\cdot)$, $\Omega(\cdot)$, $\omega(\cdot)$ and $\Theta(\cdot)$ notation. The use of a tilde, such as $\tilde{O}(\cdot)$, will indicate a polylogarithmic overhead, i.e. $\tilde{O}(f) = O(f \text{poly} \log f)$. We use $f = \text{poly}(n)$ to mean that there is a fixed polynomial p such that $f(n) \leq p(n)$ for all n .

3 Viable sets, the merge process, and area laws

Recall that our goal is to formalize and analyze an RG-like transformation in the spirit of the following claim, which for ease of explanation we state for the gapped degenerate case:

Proposition 3.1. Let H be a local Hamiltonian satisfying Assumption (DG) (q.v. Definition 2.1).

1. For every length scale ℓ and contiguous block A of ℓ qudits, there is a subspace $S \subseteq \mathcal{H}_A$ of dimension $q = r^{1+o(1)} e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$ such that S approximates the ground space T of H , in the following sense: every state in T has large overlap with a state whose reduced density on \mathcal{H}_A is supported on S .

2. Suppose given two subspaces $S_1 \subseteq \mathcal{H}_1$ and $S_2 \subseteq \mathcal{H}_2$ on adjacent blocks of ℓ qudits each, such that each of S_1, S_2 has dimension at most q and approximates T . Then it is possible to generate a subspace $S \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$ of the composite system that has the same dimension q and approximates T to the same extent as S_1, S_2 .

The key feature of the second item in the proposition is that the dimension of the merged set S has not increased: the set has the same size as S_1, S_2 separately, and yet it combines all the information each of these sets holds about the restriction of the ground space T to \mathcal{H}_1 and \mathcal{H}_2 respectively.

As we will see, the first item in the proposition leads naturally to an area law and succinct MPS representations for good approximations to states in T . The proof of the first item will be obtained by iteratively performing the merging procedure described in the second item. With additional work the merging procedure can be made efficient, leading to an efficient algorithm for computing these succinct representations.

3.1 Viable sets

We formalize the notion of a subspace *approximating* another as follows.

Definition 3.2. A subspace T is δ -close to a subspace T' if

$$P_{T'} P_T P_{T'} \geq (1 - \delta) P_{T'}.$$

We say that T and T' are mutually δ -close if each is δ -close to the other, and denote by $\angle_m(T, T')$ the smallest δ such that T, T' are mutually δ -close.

Geometrically,

$$\angle_m(T, T') = 1 - \min_{\substack{x \in T \\ \|x\|=1}} \max_{\substack{x' \in T' \\ \|x'\|=1}} |x \cdot x'|^2$$

is the squared sine of the largest principal angle between the subspaces T and T' (where the cosines of the principal angles are given by the singular values of $P_T P_{T'}$); in particular the statement that T is δ -close to T' is equivalent to the fact that for every $|\psi\rangle \in T'$ there exists $|\phi\rangle \in T$ such that $|\langle\psi|\phi\rangle|^2 \geq 1 - \delta$. Note that mutually close subspaces always have the same dimension.

With a view towards working with subsystems, we extend the notion of closeness to capture approximation by subspaces defined only on one half of a factored Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$.

Definition 3.3. Given a subspace $T \subseteq \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, a subspace $S \subseteq \mathcal{H}_A$ is δ -viable for T if

$$P_T P_{S_{\text{ext}}} P_T \geq (1 - \delta) P_T, \tag{1}$$

where $S_{\text{ext}} := S \otimes \mathcal{H}_B$.

This definition generalizes the definition of a viable set from Ref. [LVV15], which was specialized to the case where T is a one-dimensional subspace containing a unique ground state. Informally it captures the notion that a reasonable approximation of T can be made using the subspace $S \otimes \mathcal{H}_B$. With the definition in place, we can make the statement of item 1 of Proposition 3.1 precise. For ease of exposition we relax the dependence of q on r to $\tilde{O}(r^2)$. This dependence can be improved from $\tilde{O}(r^2)$ to $r^{1+o(1)}$ as sketched at the end of Appendix B.

Theorem 3.4. Let H be a local Hamiltonian satisfying Assumption (DG). Then for any block $A \subseteq \{1, \dots, n\}$ of $\ell \leq n$ qudits there exists a .015-viable set $S \subset \mathcal{H}_A$ for the ground space T of H of dimension at most $q = \tilde{O}(r^2) e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$.

We further note that the dependence of q on the dimension r can be improved to $\tilde{O}(r)$, using a direct “bootstrapping” argument that is slightly different from the more “algorithmic” argument that we give in Section 3.3. This improved bound and bootstrapping argument are presented in Section 7.

While the notion of viable set is quite intuitive for small δ , our arguments will also involve viable sets with parameter δ close to 1, a regime where there is less intuition. A helpful interpretation of the definition is that it formalizes the fact that for a viable set S , the image of the unit ball of S_{ext} when projected to T contains the ball of radius $(1 - \delta)$.

Lemma 3.5. *If S is δ -viable for T then for every $|t\rangle \in T$ of unit norm, there exists an $|s\rangle \in S_{ext}$ such that $P_T|s\rangle = |t\rangle$ and $\|s\| \leq \frac{1}{1-\delta}$.*

The proof of this and the following two useful lemmas appear in Appendix A.1. The first summarizes the effect of tensoring two viable sets supported on disjoint spaces.

Lemma 3.6. *Suppose S_1, S_2 are δ_1 -viable and δ_2 -viable for T respectively, defined on disjoint sets of qudits. Then the set $S := S_1 \otimes S_2$ is $(\delta_1 + \delta_2)$ -viable for T .*

The second lemma shows that our notion of closeness can be chained together and is compatible with the notion of viable set:

Lemma 3.7. *If T is δ -close to T' and T' is δ' -close to T'' then T is $2(\delta + \delta')$ -close to T'' . Consequently if S is δ -viable for T and T is δ' -close to T' then S is $2(\delta + \delta')$ -viable for T' .*

3.2 The Merge Process

We are ready to outline the merging procedure referred to in item 2 of Proposition 3.1, which lies at the heart of our RG transformation. Assume we are given a decomposition $\mathcal{H} = \mathcal{H}_L \otimes (\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_R$ of the global Hilbert space. The merge process **Merge** takes as input two subsets $V_1 \subseteq \mathcal{H}_1$ and $V_2 \subseteq \mathcal{H}_2$ and returns a subset $V \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$. To do so, it requires two additional inputs: a finite set of operators $\{A_i\}_{i=1}^{D^2}$ each acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$, along with a positive integer s . The procedure consists of the following three simple steps.

Merge($V_1, V_2, \{A_i\}, s$):

Step 1: Tensoring. Set $W = V_1 \otimes V_2$.

Step 2: Random Sampling. Let $W' \subseteq W$ be a random s -dimensional subspace of W .

Step 3: Error Reduction. Set $V = \text{Span}(\cup_i A_i W')$.

Return V .

The effectiveness of **Merge** relies on the properties of the operators $\{A_i\}$, with a sufficiently good choice of these operators leading to a formalization of item 2. of Proposition 3.1. In Section 6 we show how suitable operators can be obtained from the decomposition of an *approximate ground state projection* (AGSP). The following theorem summarizes the essential properties of the resulting $\{A_i\}$. Its proof is given in Section 6.4.1.

Theorem 3.8 (Existence of AGSP, (DG)). *Let H be a local Hamiltonian satisfying Assumption (DG), and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the n -qudit space in three contiguous blocks. There exists a collection of D^2 operators $\{A_i\}_{i=1}^{D^2}$ acting on \mathcal{H}_M along with a subspace $\tilde{T} \subseteq \mathcal{H}$ such that:*

- $\angle_m(T, \tilde{T}) \leq .005$,
- $D = e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$,
- There is $\Delta > 0$ such that $D^{12}\Delta \leq \frac{1}{2000}$ and whenever $S \subseteq \mathcal{H}_M$ is δ -viable for \tilde{T} then $S' = \text{Span}\{\cup_i A_i S\}$ is δ' -viable for \tilde{T} , with $\delta' = \frac{\Delta}{(1-\delta)^2}$.

Given a finite collection of operators $\{A_i\}$ we denote by $\{A_i\}^k$ the set of all products of k of the A_i . The following theorem states the guarantees offered by the **Merge** process when initialized with operators $\{A_i\}$ satisfying the guarantees of Theorem 3.8.

Theorem 3.9. *Let H be a local Hamiltonian satisfying Assumption (DG), and $\mathcal{H} = \mathcal{H}_L \otimes (\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_R$ a decomposition of the n -qudit space into contiguous blocks. Let $\{A_i\}$ and D be as in Theorem 3.8,*

$$s \geq 1600r(\log r + 1) \quad \text{and} \quad k = \frac{1}{2} \lceil \log_D(s) \rceil.$$

Let $V_1 \subseteq \mathcal{H}_1$ and $V_2 \subseteq \mathcal{H}_2$ be .015-viable subspaces for T of size $q = s^2$ each. Then with probability $1 - e^{-\Omega(s)}$ the space $V = \text{Merge}(V_1, V_2, \{A_i\}^k, s)$ is .015-viable for T with $|V| \leq q$.

The proof below analyzes the effect of each of the three steps of the **Merge** process. The first creates the trivial subspace $V_1 \otimes V_2$, whose dimension $q^2 = \dim(V_1) \dim(V_2)$ is too large, and whose overlap with T is worse than desired by a factor of 2. The random sampling step roughly evenly trades off size for overlap: it picks a random s dimensional subspace for $s \ll q$, at the expense of making the overlap roughly $\frac{s}{q^2}$. Finally, the application of the AGSP (via the operators $\{A_i\}$) blows up the size from s to at most q , while increasing overlap to at least the original overlap of V_1 and V_2 . This relies on the highly favorable D, Δ -tradeoff of the AGSP.

Proof. We analyze each of the three steps of the **Merge** process:

1. *Tensoring.* Applying Lemma 3.6 yields that the result of step 1, $W = V_1 \otimes V_2 \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$, is a .03 viable set for T of size q^2 . Using the first condition from Theorem 3.8 and applying Lemma 3.7, W is .07-viable for \tilde{T} .

2. *Random Sampling.* We show that at the end of this step, with high probability W' is $(1 - \alpha)$ -viable for \tilde{T} with $\alpha = (.8)s/q^2$. We accomplish this by establishing that with high probability $\|P_{W'_{ext}}|v\rangle\|^2 \geq \alpha$ for all states $|v\rangle \in \tilde{T}$, where $W'_{ext} = \mathcal{H}_L \otimes W' \otimes \mathcal{H}_R$ and $W_{ext} = \mathcal{H}_L \otimes W \otimes \mathcal{H}_R$.

Let $|v\rangle \in \tilde{T}$ have norm 1, and $|w\rangle = P_{W_{ext}}|v\rangle \in W_{ext}$. Using that W is .07-viable for \tilde{T} it follows that $\| |w\rangle \|^2 \geq .995$. Since $W'_{ext} \subseteq W_{ext}$, $P_{W'_{ext}}|v\rangle = P_{W'_{ext}}|w\rangle$. Applying a standard concentration argument (q.v. Corollary A.3 with $\varepsilon = .1$) it holds that $\|P_{W'_{ext}}|v\rangle\|^2 \geq (.9)(.995)\frac{s}{q^2}$ with probability at least $1 - q^2 e^{-s/400}$.

By a simple volume argument (see e.g. [Ver10, Lemma 5.2]) there exists a $\nu = \sqrt{(.1)(.9)(.995)\frac{s}{q^2}}$ -net for the Euclidean unit ball of \tilde{T} consisting of at most $(1 + \frac{2}{\nu})^r$ elements of \tilde{T} . Applying the preceding argument to each $|v\rangle$ in the net, a choice of s such that

$$\eta = \left(1 + \frac{2}{\nu}\right)^r q^2 e^{-s/400} < 1 \tag{2}$$

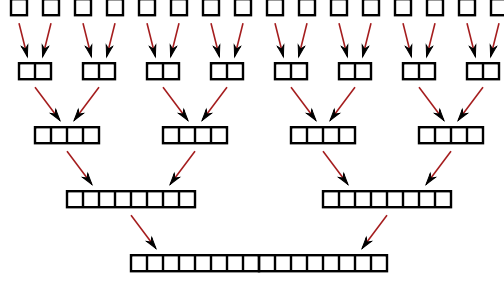


Figure 1: The parallel structure of Procedure 1. Each square represents a qudit, and successive viable sets are supported on neighboring groups of squares.

will guarantee that with probability at least $1 - \eta$, $\|P_{W'_{ext}}|v\rangle\|^2 \geq (.9)(.995)\frac{s}{q^2}$ for all $|v\rangle$ in the net; hence $\|P_{W'_{ext}}|v\rangle\|^2 \geq (.99)(.9)(.995)\frac{s}{q^2} \geq .8\frac{s}{q^2}$ for all $|v\rangle \in \tilde{T}$ of unit norm. The equation (2) is satisfied with

$$s > 400 \left(2 \log q + \frac{r}{2} \log \left(1 + \sqrt{47 \frac{q^2}{s}} \right) \right),$$

a condition verified by the choices of s and q made in the theorem.

Step 3: Error Reduction. Applying Theorem 3.8 k times in sequence, $V = \text{Span}\{\{A_i\}^k \cdot W'\}$ is $\frac{\Delta^k}{(1-\delta)^2} = \frac{\Delta^k}{\alpha^2}$ -viable for T' of size at most $D^{2k}s$. Our choice of k ensures $D^{2k}s = q$, and the relation between D and Δ implies that

$$\frac{\Delta^k}{\alpha^2} = \frac{s^6 \Delta^k}{.64} = \frac{D^{12k} \Delta^k}{.64} \leq \frac{1}{(.64)} \frac{1}{2000} \leq .001.$$

Thus V is .001-viable for \tilde{T} , and by Lemma 3.7 it is $.012 < .015$ -viable for T . \square

3.3 Area law for degenerate Hamiltonians

In this section we first prove Theorem 3.4 establishing the claim made in the first item of Proposition 3.1. From the theorem we then deduce an area law for local Hamiltonians satisfying Assumption (DG) (degenerate ground space with a spectral gap; q.v. Definition 2.1).

Proof of Theorem 3.4. Consider a system A of $\ell \leq n$ consecutive qudits; for ease of notation we'll assume that ℓ is a power of 2 and A consists of the first ℓ qudits of the n -qudit chain on which H acts. The proof of the theorem is based on the following iterative procedure for constructing the .015-viable set claimed in the theorem. The procedure depends on a set of operators $\{A_i\}$ obtained from Theorem 3.8 for various decompositions of \mathcal{H} , and we let s and k be as in the theorem.

Procedure 1. Given a local Hamiltonian H satisfying (DG), returns a viable set for T supported on the first ℓ qudits.

Initialization. Set $V_j^0 = \mathcal{H}_j$ for $j = 1, 2, \dots, \ell$.

Iteration. For $i = 1, \dots, \log(\ell)$ do:

For all $j \in \{1, 2, \dots, \frac{\ell}{2^i}\}$, set

$$V_j^i = \mathbf{Merge}(V_{2j-1}^{i-1}, V_{2j}^{i-1}, \{A_i\}^k, s) \subseteq \mathcal{H}_{[(j-1)2^i+1, j2^i]},$$

where $\{A_i\}$ are as in Theorem 3.8 for the decomposition $\mathcal{H} = \mathcal{H}_{[1, (j-1)2^i]} \otimes (\mathcal{H}_{[(j-1)2^i+1, (2j-1)2^i-1]} \otimes \mathcal{H}_{[(2j-1)2^i+1, j2^i]}) \otimes \mathcal{H}_{[j2^i+1, \ell]}$, and s and k are as in Theorem 3.9.

The proof of Theorem 3.4 follows by showing that with positive probability Procedure 1 returns a subspace $V_1^{\log \ell}$ that is .015-viable for T and such that $|V_1^{\log \ell}| = \tilde{O}(r^2) e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$. Let $q = D^{2k} s$ be the size of the space output by **Merge**, and observe that by Theorem 3.8 it holds that $q = \tilde{O}(r^2) e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$. We prove the result by induction, showing that V_j^i is .015-viable for T with $|V_j^i| \leq q$ for each i, j . The initialization step establishes this for $i = 0$ since each V_j^0 is 0-viable for T with $|V_j^0| = d$. The induction step is a direct consequence of Theorem 3.9, which establishes that at each iteration with probability $1 - e^{-\Omega(s)}$ the set $V_j^i = \mathbf{Merge}(V_{2j-1}^{i-1}, V_{2j}^{i-1}, s, \{A_i\}^k)$ is .015-viable for T with $|V_j^i| \leq q$. Each merging operation succeeds with independent probability, therefore there is a positive probability that the procedure terminates with a .015-viable set $V_1^{\log \ell}$ for T . \square

An area law for gapped Hamiltonians with a degenerate ground space follows readily from Theorem 3.4. Indeed, for any desired cut Theorem 3.4 establishes the existence of 0.015-viable sets of size at most q for the block of qudits on either side of the cut. As a consequence each element of the ground space T has a constant approximation by a state of Schmidt rank at most q . Applying a suitable AGSP to the tensor product of two such viable sets one can obtain a δ -viable set, for any desired δ , at a modest (depending on δ) increase in size. This kind of trade-off leads to a standard proof bounding both the Schmidt rank and the von Neumann entropy across the cut for any state in the ground space. We state the result here, relegating the proof to Appendix B.

Corollary 3.10 (Area law for degenerate gapped Hamiltonians). *Let H be a local Hamiltonian satisfying Assumption (DG). For any cut and any $\delta = \text{poly}^{-1}(n)$ there is subspace $S \subseteq \mathcal{H}$ that is δ -close to T and such that every element of S has Schmidt Rank no larger than*

$$s(\delta) = \tilde{O}(r^2) e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)} \cdot e^{\tilde{O}(\gamma^{-1/4} \log^{3/4}(\frac{1}{\delta}) \log d)}.$$

Moreover, every state $|\psi\rangle \in T$ has entanglement entropy

$$S(|\psi\rangle\langle\psi|) \leq 4 \log r + \tilde{O}\left(\frac{1}{\gamma} \log^3 d\right)$$

and can be approximated by a state $|\psi'\rangle$ such that $|\langle\psi|\psi'\rangle| > 1 - \delta$ and $|\psi'\rangle$ has an MPS representation with bond dimension bounded by

$$\tilde{O}(r^2) e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)} e^{\tilde{O}(\gamma^{-1/4} \log^{3/4}(\frac{n}{\delta}) \log d)}.$$

We note that the dependence on r in the bounds for the Schmidt rank and the bond dimension of the MPS approximation can be improved from $\tilde{O}(r^2)$ to $r^{1+o(1)}$. A sketch for these improvements is given in Appendix B. We do not go through them in detail because there is a simpler way of getting these bounds through a clean “bootstrapping” argument, given in Section 7. The bootstrapping argument has the drawback of being non-constructive, whereas the approach followed here is well-tuned to developing the algorithms exposed in the next section. As shown in Section 5 it also works in the case of gapless Hamiltonians satisfying Assumption (LD).

4 Moving to algorithms

There are two main obstacles to turning Procedure 1 into an efficient algorithm. The first consists in showing that operators $\{A_i\}$ satisfying the conditions of Theorem 3.8 can be generated efficiently from a description of the Hamiltonian, and that it is possible to apply these operators efficiently, as required to complete the error reduction step of the **Merge** process. The following theorem states that this can be achieved.

Theorem 4.1 (Efficient AGSP, (DG)). *There exists a procedure **Generate** (H, M, ε'_M) which takes as input*

- *A local Hamiltonian H satisfying Assumption (DG),*
- *A decomposition $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ of the n -qudit space into contiguous blocks,*
- *An estimate ε'_M for the minimal energy ε_M of the restriction of H to \mathcal{H}_M such that $|\varepsilon_M - \varepsilon'_M| \leq 10$,*

and returns

- *MPO representations for a collection of D^2 operators $\{A_i\}_{i=1}^{D^2}$ acting on \mathcal{H}_M and of bond dimension at most $n^{\tilde{O}(\gamma^{-2})}$ satisfying the conditions of Theorem 3.8 for some subspace \tilde{T} ,*
- *An MPO for an operator \tilde{H}_M such that $\|\tilde{H}_M\| = O(\gamma^{-1} \log \gamma^{-1})$ and the minimal energy $\tilde{\varepsilon}_M$ of \tilde{H}_M restricted to \tilde{T} satisfies $|\varepsilon_M - \tilde{\varepsilon}_M| < 1/2$.*

*Moreover, **Generate** (H, M, ε'_M) runs in time $n^{\tilde{O}(\gamma^{-2})}$.³*

The proof of Theorem 4.1 relies on new constructions of approximate ground state projections (AGSP), and we give it in Section 6.5.2. The theorem guarantees that the $\{A_i\}$ can be constructed efficiently *provided* it is possible to provide a good approximation to the ground state energy of the restriction of H to \mathcal{H}_M . This is the reason for including \tilde{H}_M as parts of the output of **Generate**, which is then used by an additional step of *energy estimation* incorporated in Algorithm 1.

The second difficulty encountered in turning Procedure 1 into an algorithm is that, even if the $\{A_i\}$ can be applied efficiently, due to the logarithmic number of iterations it may be that the bond dimension of MPS representations for the elements of the viable sets we work with increase to super-polynomial. This difficulty can be overcome by introducing a *bond trimming* component Trim_ξ to the **Merge** procedure, resulting in the following modified procedure **Merge'** taking an additional trimming parameter ξ as input (ξ will usually be of order $\text{poly}^{-1}(n)$):

Merge' $(V_1, V_2, \{A_i\}, s, k, \xi)$:

Step 1: Tensoring. Set $W = V_1 \otimes V_2$.

Step 2: Random Sampling. Let $W' \subseteq W$ be a random s -dimensional subspace of W .

Step 3: Error Reduction. Set $V = W'$. Repeat k times:

Set $V = \text{Trim}_\xi(\text{Span}(\cup_i A_i V))$.

Return V .

Correctness of **Merge'** (for an appropriate choice of ξ) is based on the area law proven in Corollary 3.10. The details of the trimming⁴ procedure Trim_ξ , together with the analysis of **Merge'**, are given in Appendix C. We describe the resulting algorithm.

³Here and in all our estimates on running times we suppress dependence on the local dimension d , which is treated as a constant.

⁴We note that the trimming procedure differs from that of [LVV15]

Algorithm 1. Given a local Hamiltonian H satisfying Assumption (GS), returns a set $V_1^{\log n}$ that is 0.015-close to the ground space T of H .

Initialization. Set $V_j^0 = \mathcal{H}_j$ and $\varepsilon'_{0,j} = 0$ for $j \in \{1, 2, \dots, n\}$.

Iteration. For $i = 1, \dots, \log(n)$ and $j \in \{1 \dots \frac{n}{2^i}\}$ do:

Generate. Let $M = \{(j-1)2^i, (j-1)2^i + 1, \dots, j2^i - 1\}$, $\varepsilon'_M = \varepsilon'_{i-1,2j-1} + \varepsilon'_{i-1,2j}$. Set $(\{A_i\}, \tilde{H}_M) = \mathbf{Generate}(H, M, \varepsilon'_M)$.

Merge. Set $V_j^i = \mathbf{Merge}'(V_{2j-1}^{i-1}, V_{2j}^{i-1}, \{A_i\}, s, k, \xi) \subseteq \mathcal{H}_{[(j-1)2^i+1, j2^i]}$, where s and k are specified in Theorem 3.9 and $\xi = \text{poly}^{-1}(n, r)$ is chosen small enough (see proof of Theorem 4.2).

New Energy Estimation. Form the subspace $V = \{A_i\}^t \cdot (V_{2j-1}^{i-1} \otimes V_{2j}^{i-1})$ for $t = \Theta(\log \gamma^{-1})$. Compute the smallest eigenvalue $\varepsilon'_{i,j}$ of the restriction of \tilde{H}_M to V .

Final step. Return $V_1^{\log n}$.

The following theorem proves the correctness of Algorithm 1.

Theorem 4.2. Let H be a local Hamiltonian satisfying Assumption (DG). Then with probability at least $1 - \frac{1}{n}$ the set $V_1^{\log n}$ returned by Algorithm 1 is 0.015-viable for T .⁵ The running time of the algorithm is $n^{\tilde{O}(\gamma^{-2})}$.

Proof. The proof mirrors the analysis of Procedure 1 given in the proof of Theorem 3.4 in Section 3.3; the two main differences are that we must show that at every step, with high enough probability the call to **Merge'** yields a good viable set and the **New Energy Estimation** step yields a sufficiently accurate energy estimate for the next iteration.

Both conditions are satisfied at initialization since each V_j^0 is 0-viable for T with $|V_j^0| = d$ and the energy estimates are accurate since there are no terms of the Hamiltonian when restricting to single particles.

Assume V_{2j-1}^{i-1} and V_{2j}^{i-1} are both .015-viable for T with $|V_{2j-1}^{i-1}|, |V_{2j}^{i-1}| \leq q$, and $\varepsilon'_{i-1,2j-1}, \varepsilon'_{i-1,2j}$ both within an additive ± 3 of their respective true values (the ground state energy of the restriction of H to the corresponding spaces). As a result ε'_M is within 7 of the correct value ε_M , and by Theorem 4.1 **Generate** yields a set $\{A_i\}$ with the properties stated in Theorem 3.8. Thus by Theorem 3.9 V_j^i is .015-viable for T with probability $1 - e^{-\Omega(s)} \geq 1 - \frac{1}{n^2}$ (provided $r \geq \log n$, which we may always assume without loss of generality). For this we need to check Theorem 3.9 still applies when **Merge** is replaced by **Merge'**. The analysis of the trimming procedure given in Lemma C.2 shows that this is the case provided the error reduction parameter Δ associated with the $\{A_i\}$ is replaced by $(\Delta + 2\sqrt{krs\xi})$; choosing $\xi = \text{poly}^{-1}(n, r)$ we may ensure that $2\sqrt{krs\xi} < .0001D^{-12}$ (using Corollary 3.10 to bound s by a polynomial at the expense of replacing T by a set that is $\text{poly}^{-1}(n)$ -close to T). With this choice, the remaining calculation of 3.9 applies to still yields that V_j^i is .015-viable for T .

Once this has been established, an application of the third item from Theorem 3.8 shows that provided the constant implicit in the definition of t is chosen large enough the subspace V obtained after the **New Energy Estimation** step is $O(\gamma^2)$ -viable for \tilde{T} . Using that $\|\tilde{H}_M\| = O(\gamma^{-1} \log \gamma^{-1})$ it follows that $\varepsilon'_{i,j}$ is

⁵The probability of success can be improved to $1 - \text{poly}^{-1}(n)$ by scaling the parameter s used in the algorithm by an appropriate constant.

within an arbitrarily small constant of the minimal energy of \tilde{H}_M restricted to \tilde{T} . Using the last guarantee from Theorem 4.1, $\varepsilon'_{i,j}$ is within $\frac{3}{2}$ of the minimal energy ε_M of the restriction of H to \mathcal{H}_M . This completes the inductive step.

We have shown that the iterative step succeeds with probability at least $1 - 1/n^2$; since there are a total of n such merging steps, applying a union bound the final set $V_1^{\log n}$ is .015-viable with probability at least $1 - \frac{1}{n}$.

In total the complete algorithm requires only a polynomial number of operations on MPS representations of vectors. Due to trimming, all these vectors have polynomial bond dimension and thus each operation can be implemented in polynomial time. The complexity is dominated by the complexity of the procedure **Generate** and the application of the operators A_i , which is $n^{\tilde{O}(\gamma^{-2})}$. \square

We end this section by noting that in case one desires a better than constant approximation to T the final step of Algorithm 1 can be replaced by the following:

Final step. Set $K = (\mathbb{1} - H/\|H\|)$ and $\tau = 10\|H\|\gamma^{-1}\log(1/\delta)$. Choose an orthonormal basis $\{|y_i^{(0)}\rangle\}$ for $V_1^{\log n}$. Repeat for $t = 1, \dots, \tau$:

$$\text{Set } \{|y_i^{(t)}\rangle\} = \text{Trim}_\xi(\text{Span}\{K|y_i^{(t-1)}\rangle\}).$$

Return $\{|z_i\rangle\}$, the smallest r eigenvectors of H restricted to $W = \text{Span}\{|y_i^{(\tau)}\rangle\}$.

We claim that the result of this step is a basis $\{|z_i\rangle\}$ for a subspace S such that $\angle_m(S, T) \leq \delta$. Recall from Theorem 4.2 that we have the guarantee that the set $V_1^{\log n}$ is 0.015-viable for T . Let $|\psi\rangle$ be an eigenvector of H with eigenvalue ε_0 , and $|v\rangle \in V_1^{\log n}$ such that $|v\rangle = \alpha|\psi\rangle + \sqrt{1 - |\alpha|^2}|v^\perp\rangle$, where $\alpha \geq 0.9$ and $|v^\perp\rangle$ is supported on eigenvectors of H with eigenvalue at least $\varepsilon + \gamma$. Following the same analysis as given in the proof of Lemma 6.2 (error reduction for AGSP) it follows that after renormalization the overlap of $K|v\rangle/\|K|v\rangle\|$ with $|v\rangle$ has improved from α to

$$\frac{\alpha^2}{\alpha^2 + (1 - \alpha^2)(1 - \gamma/\|H\|)} = \frac{\alpha^2}{1 - \gamma(1 - \alpha^2)/\|H\|} \geq \alpha^2 \left(1 + \frac{\gamma}{2\|H\|}\right).$$

Thus the set $K\{|y_i^{(1)}\rangle\}$ is $0.9(1 + \gamma/(2\|H\|))$ -viable for T . Assuming ξ is chosen small enough, by Lemma C.2 the set $\{|y_i^{(2)}\rangle\}$ will remain $0.9(1 + \gamma/(3\|H\|))$ -viable for T . Repeating this procedure τ times yields a set W that is δ -viable for T . Finally, each of the r vectors $|z_i\rangle$ returned by the algorithm must have energy at most $\varepsilon_0 + \delta\gamma$, which using the spectral gap condition implies that $\text{Span}\{|z_i\rangle\}$ and T are mutually δ -close.

4.1 Frustration-free Hamiltonians with a unique ground state

The computation-intensive step of the AGSP-based RG transformation introduced in Section 3 is the construction and subsequent application of the set of operators $\{A_i\}$. In the special case where the Hamiltonian H satisfies Assumption (FF), i.e. H is frustration-free and has a spectral gap, the operators $\{A_i\}$ can be constructed very efficiently, yielding strong bounds on the running time. We state a specialized theorem for this setting. The proof is given in Section 6.3.

Theorem 4.3 (Efficient AGSP, (FF)). *Let H be a local Hamiltonian satisfying Assumption (FF), and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the n -qudit space into contiguous regions. There exists a procedure **Generate 2**(H, M) which takes as input*

- A local Hamiltonian H satisfying Assumption (DG),
- A decomposition $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ of the n -qudit space into contiguous blocks,

and returns MPO representations for a collection of D^2 operators $\{A_i\}_{i=1}^{D^2}$ acting on \mathcal{H}_M such that the following hold:

- $D = 2^{\tilde{O}(\gamma^{-1} \log^3 d)}$
- There is $\Delta > 0$ such that $D^{12}\Delta < \frac{1}{2000}$ and for any $S \subseteq \mathcal{H}_M$ that is δ -viable for $\{|\Gamma\rangle\}$ it holds that $S' = \text{Span}\{\cup_i A_i S\}$ is δ' -viable for T with $\delta' = \frac{\Delta}{(1-\delta)^2}$.
- Each A_i has bond dimension at most $2^{\tilde{O}(\gamma^{-2} \log^5 d)}$.

Moreover, for constant d and $\gamma > 0$ the procedure **Generate 2**(H, M, ε'_M) runs in time $n^{(1+o(1))}$.

We note that in the case where M consists of all n qudits the procedure returns a single operator A acting on the whole space. The following algorithm is an adaptation of Algorithm 1 to the case of frustration-free Hamiltonians:

Algorithm 2. Given a local Hamiltonian H satisfying Assumption (FF), returns a δ -approximation to its ground state $|\Gamma\rangle$.

Initialization. Set $V_j^0 = \mathcal{H}_j$ for $j \in \{1, 2, \dots, n\}$.

Iteration. For $i = 1, \dots, \log(n)$ and all $j \in \{1, \dots, \frac{n}{2^i}\}$,

Generate. Let $M = \{(j-1)2^i, (j-1)2^i + 1, \dots, j2^i - 1\}$. Set $\{A_i\} = \mathbf{Generate\ 2}(H, M)$.

Merge. Set $V_j^i = \mathbf{Merge}'(V_{2j-1}^{i-1}, V_{2j}^{i-1}, \{A_i\}, s, k, \xi) \subseteq \mathcal{H}_{[(j-1)2^i+1, j2^i]}$, where k, s are as in Theorem 3.9 (with $r = 1$) and $\xi = \tilde{\Theta}(n^{-1/2})$.

Final step. Let K be the unique operator A computed at the last iteration, and $\tau = 10\|H\|\gamma^{-1} \log(1/\delta)$. Choose an orthonormal basis $\{|y_i^{(0)}\rangle\}$ for $V_1^{\log n}$. Repeat for $t = 1, \dots, \tau$:

Set $\{|y_i^{(t)}\rangle\} = \text{Trim}_\xi(\text{Span}\{|Ky_i^{(t-1)}\rangle\})$, for $\xi = \tilde{\Theta}(n^{-1/2})$.

Return the smallest eigenvector $|z\rangle$ of H restricted to $W = \text{Span}\{|y_i^{(\tau)}\rangle\}$.

Theorem 4.4. Let H be a local Hamiltonian satisfying Assumption (FF) and $\delta = n^{-\omega(1)}$. With probability at least $1 - \frac{1}{n}$ the vector $|z\rangle$ returned by Algorithm 3 is such that $|\langle z|\Gamma\rangle| \geq 1 - \delta$. Moreover the algorithm runs in time $O(n^{1+o(1)}M(n))$, where $M(n) = O(n^{2.38})$ denotes matrix multiplication time.

Proof. The proof follows the same outline as that of Theorem 4.2 analyzing Algorithm 1. To ensure the algorithm is efficient, it is important to choose the trimming parameter ξ to be as large as possible. It follows from the area law for 1D gapped systems [AKLV13] that the ground state $|\Gamma\rangle$ of H can be approximated up to accuracy $\text{poly}^{-1}(n)$ by a matrix product state with sub-linear bond dimension. Thus by Lemma C.2, using that r, s are both constant a choice of $\xi = n^{-(1/2+\omega(1))}$ will suffice to ensure the error remains negligible, while also maintaining the property that all MPS manipulated have essentially linear bond dimension. The

essential operations on such vectors required in the algorithm, such as multiplication by an MPO A_i of constant bond dimension, or writing in canonical form, can all be computed in time $\tilde{O}(nM(B))$ where $M(B)$ is matrix multiplication time for $B \times B$ matrices and B is an upper bound on the bond dimension of the MPS being manipulated; $M(B)$ corresponds to the cost of performing individual singular value decompositions on the tensors that form each of the MPS. The claim on the running time follows since the number of iterations of the algorithm is logarithmic. \square

5 Gapless Hamiltonians

In this section we extend the results of Section 3 and Section 4, which respectively derived an area law and efficient algorithm for local Hamiltonians satisfying the Degenerate Gapped (DG) assumption, to the case of gapless Hamiltonians satisfying the Low Density (LD) assumption described in Definition 2.1.

Both the proof of the area law, and the algorithm, follow the same outline as those introduced in the previous sections. The main obstacle, of course, consists in dealing with a gapless system. Recall that the existence of a spectral gap seemed crucial for the construction of the operators $\{A_i\}$ that underlie both results.

What makes it possible to tackle the gapless case are the strong properties of a viable set. Suppose that S is a viable set for T , the set of states of energy at most η . Then S is also a viable set for T' , the set of states of energy at most $\eta - \mu$ for an arbitrary choice of μ . Now, if we apply an AGSP which amplifies the norm of states with energy less than $\eta - \mu$, and decreases the norm of states with energy greater than η , this is guaranteed to improve the quality of the viable set. This is because by Lemma 3.5, for each state in T' the viable set S contains an approximation to that state that is guaranteed to have no projection onto the orthogonal complement of T' in T . In this sense, regarding S as a viable set for T' creates a virtual spectral gap $\mu > 0$. This approach is formalized below.

5.1 Area Law

The following two theorems are direct analogues of Theorem 3.8 and Theorem 4.1 respectively. The first theorem guarantees the existence of operators $\{A_i\}$ with good properties. The proof is given in Section 6.4.1.

Theorem 5.1 (Existence of AGSP, (LD)). *Let $\mu > 0$ be a constant, H a local Hamiltonian satisfying Assumption (LD), and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the n -qudit space in three contiguous blocks. For any $\eta \geq \eta_1 \geq 2\frac{\mu}{\log n}$ there exists a collection of D^2 operators $\{A_i\}_{i=1}^{D^2}$ acting on \mathcal{H}_M along with two subspaces $\tilde{T}_- \subseteq \tilde{T} \subseteq \mathcal{H}$ such that:*

- $H_{[\varepsilon_0, \varepsilon_0 + \eta_1]}$ is .005-close to \tilde{T} ,
- \tilde{T}_- is .005-close to $H_{[\varepsilon_0, \varepsilon_0 + \eta_1 - \frac{\mu}{\log n}]}$,
- $D = e^{\tilde{O}\left(\frac{\log n}{\mu} \log^3 d\right)}$,
- There is a $\Delta > 0$ such that $D^{12}\Delta < \frac{1}{2000}$ and for any $k \geq 1$ and $S \subseteq \mathcal{H}_M$ that is δ -viable for \tilde{T} it holds that $S' = \text{Span}\{\cup_{A \in \{A_i\}^k} AS\}$ is δ' -viable for \tilde{T}_- with $\delta' = \frac{\Delta^k}{(1-\delta)^2}$.

The second theorem states the properties of the **Merge** procedure, when initialized with operators $\{A_i\}$ satisfying the conditions of the previous theorem.

Theorem 5.2. Let $\mu < \eta$ be a positive constant, H a local Hamiltonian satisfying Assumption (LD), and $\mathcal{H} = \mathcal{H}_L \otimes (\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_R$ a decomposition of the n -qudit space in contiguous blocks. Let $\{A_i\}$ and D be as in Theorem 5.1,

$$s \geq 1600r(\log r + 1) \quad \text{and} \quad k = \frac{1}{2} \lceil \log_D(s) \rceil.$$

Let $V_1 \subseteq \mathcal{H}_1$ and $V_2 \subseteq \mathcal{H}_2$ be .015-viable subspaces for $T = H_{[\varepsilon_0, \varepsilon_0 + \eta]}$ of size $q = s^2$ each. Then with probability $e^{-\Omega(s)}$ the space $V = \text{Merge}(V_1, V_2, \{A_i\}^k, s)$ is .015-viable for $T_- = H_{[\varepsilon_0, \varepsilon_0 + \eta - \frac{\mu}{\log n}]}$ with $|V| \leq q$.

Proof. The proof follows the same outline as that of Theorem 3.9. The first two steps (tensoring, random sampling) are identical, and guarantee that we initiate the error reduction step with a set W' that is $(1 - \alpha)$ -viable for \tilde{T} , with $\alpha = .8 \frac{s}{q^2}$. The analysis of that step is also similar as before, except we use the fourth item from Theorem 5.1 and replace \tilde{T} and T in the proof of Theorem 3.9 with \tilde{T}_- and T_- here. \square

With Theorem 5.2 in place, the path towards an area law in the gapless case follows the same as in Section 3.3. The following is a direct adaptation of Procedure 1.

Procedure 2. Given a local Hamiltonian H satisfying (LD), returns a viable set for T supported on the first ℓ qudits.

Initialization. Let $\mu < \eta$ be a positive constant. Set $V_j^0 = \mathcal{H}_j$ for $j \in \{1, 2, \dots, \ell\}$.

Iteration. For $i = 1, \dots, \log(\ell)$ and all $j \in \{1, 2, \dots, \frac{\ell}{2^i}\}$:

Set $V_j^i = \text{Merge}(V_{2j-1}^{i-1}, V_{2j}^{i-1}, \{A_i\}^k, s) \subseteq \mathcal{H}_{[(j-1)2^i+1, j2^i]}$, where $\{A_i\}$ are as in Theorem 5.1 for the decomposition $\mathcal{H} = \mathcal{H}_{[1, (j-1)2^i]} \otimes \mathcal{H}_{[(j-1)2^i+1, j2^i]} \otimes \mathcal{H}_{[j2^i+1, \ell]}$ and $\eta_1 = \eta - \frac{(i-1)\mu}{\log n}$, and s and k are as in Theorem 5.2.

Theorem 5.3. Let H be a local Hamiltonian satisfying Assumption (LD). For any block $A \subseteq \{1, \dots, n\}$ of $\ell \leq n$ qudits the subspace $V_1^{\log \ell}$ returned by Procedure 2 is .015-viable for $H_{[\varepsilon_0, \varepsilon_0 + \eta - \mu]}$ and has size $|V_1^{\log \ell}| = \tilde{O}(r^2) e^{\tilde{O}(\frac{\log n}{\mu} \log^3 d)}$.

Proof. The proof mirrors that of Theorem 3.4 with the role of Theorem 3.8 replaced by Theorem 5.1 and Theorem 3.9 replaced by Theorem 5.2. \square

The proof of the following result mirrors that of Corollary 3.10 and we omit it.

Corollary 5.4 (Area law for low-density Hamiltonians). Let H be a local Hamiltonian satisfying Assumption (LD), $\mu < \eta$ any positive constant and $T = H_{[\varepsilon_0, \varepsilon_0 + \eta - \mu]}$. For any cut and any $\delta = \text{poly}^{-1}(n)$ there is a subspace $S \subseteq \mathcal{H}$ that is δ -close to T and such that every element of S has Schmidt Rank no larger than $\tilde{O}(r^2) e^{\tilde{O}(\frac{\log n}{\mu} \log^3 d)} \cdot e^{\tilde{O}((\frac{\log n}{\mu})^{1/4} \log^{3/4}(\frac{1}{\delta}) \log d)}$. Moreover, every state $|\psi\rangle \in T$ has entanglement entropy

$$S(|\psi\rangle\langle\psi|) \leq 4 \log r + \tilde{O}\left(\frac{\log n}{\mu} \log^3 d\right)$$

and can be approximated by a state $|\psi'\rangle$ such that $|\langle\psi|\psi'\rangle| \geq 1 - \delta$ and $|\psi'\rangle$ has an MPS representation with bond dimension bounded by

$$\tilde{O}(r^2) e^{\tilde{O}(\frac{\log n}{\mu} \log^3 d)} e^{\tilde{O}((\frac{\log n}{\mu})^{1/4} \log^{3/4}(\frac{n}{\delta}) \log d)}.$$

5.2 A quasi-polynomial time algorithm

As in the case of Hamiltonians satisfying Assumption (DG) we are able to show that given a good enough approximation to the minimal energy of H restricted to \mathcal{H}_M one can generate operators $\{A_i\}$ satisfying the conditions of Theorem 5.1. Due to the absence of a constant spectral gap, and our introduction of an “artificial” gap of order $1/\log n$, the procedure now runs in quasipolynomial time $e^{\tilde{O}(\log^3 n)}$. The theorem is proved in Section 6.5.2

Theorem 5.5 (Efficient AGSP, (LD)). *There exists a procedure **Generate**($H, M, \varepsilon'_M, \eta_1, \mu$) which takes as input*

- A local Hamiltonian H satisfying Assumption (LD), a parameter $\eta_1 \leq \eta$ and a constant $\mu > 0$,
- A decomposition $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ of the n -qudit space into contiguous blocks,
- An estimate ε'_M for the minimal energy ε_M of the restriction of H to \mathcal{H}_M such that $|\varepsilon_M - \varepsilon'_M| \leq 10$,

and returns

- MPO representations for a collection of D^2 operators $\{A_i\}_{i=1}^{D^2}$ acting on \mathcal{H}_M and of bond dimension at most $e^{\tilde{O}(\log^3 n)}$ each along with set \tilde{T}_-, \tilde{T} satisfying the conditions of Theorem 5.1,
- An MPO for an operator \tilde{H}_M such that $\|\tilde{H}_M\| = \tilde{O}(\log(n/\mu))$ and the minimal energy $\tilde{\varepsilon}_M$ of \tilde{H}_M restricted to \tilde{T}_- satisfies $|\varepsilon_M - \tilde{\varepsilon}_M| < 1/2$.

Moreover, **Generate 3**($H, M, \varepsilon'_M, \eta_1, \mu$) runs in time $e^{\tilde{O}(\log^3 n)}$.

The following algorithm is a straightforward adaptation of Algorithm 1.

Algorithm 3. Given a local Hamiltonian H satisfying Assumption (LD), $0 < \mu < \eta$ and $\delta = \text{poly}^{-1}(n)$, returns an orthonormal set $\{|z_i\rangle\}$ of $r = |H_{[\varepsilon_0, \varepsilon_0 + \eta - \mu]}|$ vectors having energy at most $\varepsilon_0 + \eta - \mu + \delta$ with respect to H .

Initialization. Set $V_j^0 = \mathcal{H}_j$, $\varepsilon'_{0,j} = 0$ for $j \in \{1, \dots, n\}$.

Iteration. For $i = 1, \dots, \log(n)$ and all $j \in \{1, \dots, \frac{n}{2^i}\}$:

Generate. Let $M = \{(j-1)2^i, (j-1)2^i + 1, \dots, j2^i - 1\}$, $\varepsilon'_M = \varepsilon'_{i-1, 2j-1} + \varepsilon'_{i-1, 2j}$. Set $\{A_i\} =$

Generate 3($H, M, \varepsilon'_M, \eta - \frac{(i-1)\mu}{\log n}, \mu$)

Merge. Set $V_j^i = \text{Merge}(V_{2j-1}^{i-1}, V_{2j}^{i-1}, \{A_i\}^k, s) \subseteq \mathcal{H}_{[(j-1)2^i+1, j2^i]}$, where s and k are as in Theorem 5.2.

New Energy Estimation. Form the subspace $V = \{A_i\}^t \cdot (V_{2j-1}^{i-1} \otimes V_{2j}^{i-1})$ for $t = \Theta(\log \gamma^{-1})$. Compute the lowest eigenvalue $\varepsilon'_{i,j}$ of the restriction of \tilde{H}_M to V .

Final step. Set $K = (\mathbb{1} - H/\|H\|)$, $\tau = 10\|H\| \log(1/\delta)$ and $\xi = e^{-\Omega(\log^2 n)}$. Choose an orthonormal basis $\{|y_i^{(0)}\rangle\}$ for $V_1^{\log n}$. Repeat for $t = 1, \dots, \tau$:

Set $\{|y_i^{(t)}\rangle\} = \text{Trim}_\xi(\text{Span}\{K|y_i^{(t-1)}\rangle\})$.

Return $\{|z_i\rangle\}$, the smallest r eigenvectors of H restricted to $W = \text{Span}\{|y_i^{(\tau)}\rangle\}$.

Theorem 5.6. *Let H be a local Hamiltonian satisfying Assumption (LD). With probability at least $1 - \frac{1}{n}$ the set $\{|z_i\rangle\}$ returned by Algorithm 3 is an orthonormal set of r states each having energy at most $\varepsilon_0 + \eta - \mu + \delta$ with respect to H . Moreover the running time of the Algorithm is $2^{\text{poly} \log(n)}$.*

Proof. The proof follows closely that of Theorem 4.2 with the following simple modifications: the use of Theorem 3.9 is replaced by Theorem 5.2, Theorem 4.1 is replaced by Theorem 5.5, there is no need to introduce **Merge** (since the final running time we are obtaining is already $n^{\text{poly} \log n}$ anyways), and finally Corollary 3.10 is replaced by Corollary 5.4; as a consequence any choice of ξ for the final step that is of order $e^{-\log^{1+\omega(1)} n}$ will suffice to guarantee that trimming induces an error that is negligible compared to $\delta = \text{poly}^{-1}(n)$.

Finally, we cannot conclude directly that the r vectors $|z_i\rangle$ returned by the algorithm are low-energy eigenstates; while it does hold that each must have energy at most $\varepsilon_0 + \eta - \mu + \delta$ (since the closest vectors to $H_{[\varepsilon_0, \varepsilon_0 + \eta - \mu]}$ in W will have this property), in the absence of a spectral gap for H the $|z_i\rangle$ may still be constituted of a mixture of low-energy eigenstates with energy slightly higher than $\varepsilon_0 + \eta - \mu + \delta$. \square

6 Approximate Ground State Projections

In this section we describe constructions of Approximate Ground State Projectors (AGSPs) from which the operators $\{A_i\}$ used in our results are derived.

We start with some definitions and key properties in Section 6.1. The setup for this whole section is a local Hamiltonian H acting on a 1D chain of n qudits, numbered from 1 to n . The associated Hilbert space \mathcal{H} will often be decomposed in two parts $\mathcal{H} = \mathcal{H}_J \otimes \mathcal{H}_{\bar{J}}$, or in three, $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$. In the first case qudits in J will be numbered $\{j_0, j_0 + 1, \dots, j_1\}$, and in the second case qudits in L, M and R will be numbered $\{1, \dots, i_1\}$, $\{i_1 + 1, \dots, i_2\}$ and $\{i_2 + 1, \dots, n\}$ respectively. As usual the restriction of the Hamiltonian to those terms acting inside a region A will be denoted by H_A .

We review the general Chebyshev-based AGSP construction from [AKLV13] in Section 6.2. In the following sections we describe three different schemes to truncate a Hamiltonian while preserving its locality and keeping its low-energy eigenspace close to the low-energy eigenspace of the original Hamiltonian. The Chebyshev construction applied to the truncated Hamiltonian forms the basis for the different kinds of spectral AGSP used in the procedures of the previous sections.

In Section 6.3 we describe the simplest truncation scheme, applicable to the case of a frustration-free Hamiltonian. In Section 6.4 we introduce the most general and effective scheme of “hard truncation”, which is not algorithmically efficient but yields the necessary results for the area law results presented in Section 3. In Section 6.5 we introduce a method of “soft truncation”, which is somewhat less effective but has the advantage that it can be made efficient, as is essential for its use in the algorithms presented in Section 4 and Section 5.2.

6.1 Definition and error reduction

AGSPs were defined in the context of proofs of the 1D area law for Hamiltonians with a unique ground state [ALV12, AKLV13], and were used in algorithms for finding the ground state of a gapped 1D system [LVV15].

Whereas in previous works an AGSP was primarily constructed to approximate the projector to a *unique* ground state, here our main focus will be on the case of a degenerate ground space and low-energy states. We therefore introduce a more general definition of an AGSP, which no longer approximates the projector to a ground state. Instead, it is a local operator that increases the norms of eigenvectors in the low part of

the spectrum of H , while decreasing (*shrinking*) the norms of eigenvectors in the high energy part of the spectrum. We refer to this object as a *spectral AGSP*.

Definition 6.1 (Spectral AGSP). *Given $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ and $\eta_0 < \eta_1$, a positive semidefinite operator K on \mathcal{H} is a (D, Δ) -spectral AGSP for (H, η_0, η_1) if the following conditions hold:*

- K has a decomposition $K = \sum_{i=1}^{D^2} L_i \otimes M_i \otimes R_i$,
- H and K have the same eigenvectors,
- Eigenvalues of H smaller than η_0 correspond to eigenvalues of K that are larger than or equal to 1,
- Eigenvalues of H larger than η_1 correspond to eigenvalues of K that are smaller than $\sqrt{\Delta}$.

The advantage of an AGSP, compared to an exact projection operator, lies in the fact that one can often construct a much more *local* operator, i.e., an operator with a much smaller Schmidt rank compared to the exact projector. The existence of an AGSP of small Schmidt rank which greatly shrinks the high energy part of the spectrum can be viewed as a strong characterization of the locality properties of the ground state. A favorable scaling between these two competing aspects (in the case of unique ground states) was the key feature in recent proofs of the 1D area law [ALV12, AKLV13] via the bootstrapping lemma.

Lemma 6.2 (Error reduction — Spectral AGSP). *Let $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$, $\eta_0 < \eta_1$, and $K = \sum_{i=1}^{D^2} L_i \otimes M_i \otimes R_i$ a (D, Δ) -spectral AGSP for (H, η_0, η_1) where H has ground state energy ε_0 and has no eigenvalues in the interval (η_0, η_1) . Let $S \subseteq \mathcal{H}_M$ be a δ -viable set for $T = H_{[\varepsilon_0, \eta_1]}$ of size s .*

Then the space $V = \text{Span}\{M_i S : 1 \leq i \leq D^2\}$ has dimension no bigger than $D^2 s$ and is a δ' -viable set for $T = H_{[\varepsilon_0, \eta_0]}$ with

$$\delta' = \frac{\Delta}{(1 - \delta)^2}.$$

Proof. The bound on the dimension of V is straightforward. To show V is δ' -viable for T , begin with an arbitrary unit norm vector $|v\rangle \in T$. Set $|v'\rangle = \frac{1}{\|K^{-1}|v\rangle\|} K^{-1}|v\rangle$, where K^{-1} is the pseudo-inverse. Then $|v'\rangle$ is also an element of T . Since S is δ -viable for T , applying Lemma 3.5 there exists an $|u\rangle \in \mathcal{H}_L \otimes S \otimes \mathcal{H}_R$ whose projection onto T is, up to scaling, precisely $|v'\rangle$; thus $|u\rangle = \alpha|v'\rangle + \sqrt{1 - \alpha^2}|v^\perp\rangle$ for some $\alpha \geq 1 - \delta$ and unit $|v^\perp\rangle$ that is orthogonal to T . In particular $|v^\perp\rangle$ is supported on the span of all eigenvectors of H with eigenvalue outside of $[\varepsilon_0, \eta_1] = [\varepsilon_0, \eta_0] \cup [\eta_0, \eta_1]$ and thus by the property of K , $\|K|v^\perp\rangle\|^2 \leq \Delta$.

Applying K to $|u\rangle$ yields $K|u\rangle = \alpha'|v\rangle + K|v^\perp\rangle$ with $\alpha' = \alpha \frac{1}{\|K^{-1}|v\rangle\|} \geq \alpha$ (since $|v\rangle$ is supported on eigenvectors of K with corresponding eigenvalue at least 1). Thus

$$\begin{aligned} \left| \left\langle \frac{Ku}{\|K|u\rangle\|} \middle| v \right\rangle \right|^2 &\geq \frac{\alpha'^2}{\alpha'^2 + (1 - \alpha'^2)\Delta} \\ &\geq 1 - \frac{1}{(1 - \delta)^2} \Delta. \end{aligned}$$

□

6.2 The Chebyshev polynomial AGSP

Given a Hamiltonian H with ground energy ε_0 and a gap parameter γ , a natural way to define an approximate ground state projector is by setting $K := P_k(H)$, where P_k is a polynomial that satisfies $P_k(\varepsilon_0) = 1$ and $|P_k(x)|^2 \leq \Delta$ for every $\varepsilon_0 + \gamma \leq x \leq \|H\|$. Clearly, K preserves the ground space and reduces the norm of any eigenstate $|\phi\rangle$ of H with eigenvalue at least $\varepsilon_0 + \gamma$ as $\|K|\phi\rangle\|^2 \leq \Delta$. Moreover, the lower the degree of P_k , the lower the Schmidt rank of K at every cut. Following [AKLV13] we construct such a polynomial based on the use of Chebyshev polynomials. The construction is summarized in the following definition.

Definition 6.3 (The Chebyshev-based AGSP). *Let H be a Hamiltonian and $\eta_0 < \eta_1$ two parameters.⁶ For any integer $k > 0$, let T_k be the k -th Chebyshev polynomial of the first kind, and P_k the following rescaling of T_k :*

$$P_k(x) := \frac{1}{\tilde{P}_k(\eta_0)} \tilde{P}_k(x), \quad \text{where} \quad \tilde{P}_k(x) := T_k \left(2 \frac{x - \eta_1}{\|H\| - \eta_1} - 1 \right). \quad (3)$$

The Chebyshev AGSP of degree k for H is $K := P_k(H)$.

The properties of the Chebyshev AGSP are given in the following theorem.

Theorem 6.4. *Let H be a Hamiltonian on n qudits, $\eta_0 < \eta_1$ two parameters and $\gamma = \eta_1 - \eta_0$. Suppose that for some $u_1 < u_2 \in \{1, \dots, n\}$ and $3 \leq \ell \leq (u_2 - u_1)/2$, H can be written as*

$$H = H_L + h_{u_1-\ell} + \dots + h_{u_1} + \dots + h_{u_1+\ell-1} + H_M + h_{u_2-\ell} + \dots + h_{u_2} + \dots + h_{u_2+\ell-1} + H_R, \quad (4)$$

where each h_i is a 2-local operator on qudits $\{i, i+1\}$ and H_L , H_M and H_R are defined on qudits $J_L = \{1, \dots, u_1 - \ell\}$, $J_M = \{u_1 + \ell, \dots, u_2 - \ell\}$ and $J_R = \{u_2 + \ell, \dots, n\}$ respectively. For any integer $k > 0$ let

$$\Delta := 4e^{-4k} \sqrt{\frac{\gamma}{\|H\| - \eta_0}}.$$

The degree- k Chebyshev AGSP K for H satisfies the following properties:

1. K has the same eigenvectors as H .
2. For any any eigenvector $|\psi\rangle$ of H with eigenvalue λ such that $\lambda \leq \eta_0 + \gamma/k$, $|\psi\rangle$ is an eigenvector of K with eigenvalue $P_k(\lambda)$ satisfying

$$P_k(\lambda) \geq 1 - O\left(\frac{k|\lambda - \eta_0|}{\gamma\|H\|} \sqrt{\Delta}\right)$$

and $P_k(\lambda) \geq 1$ if $\lambda \leq \eta_0$. Moreover, any eigenvector of H with eigenvalue $\lambda = \eta_0$ is mapped to a fixed point of K .

3. For every eigenvector $|\psi\rangle$ with eigenvalue at least η_1 , $\|K|\psi\rangle\|^2 \leq \Delta$.
4. The Schmidt rank of K at all cuts in the region J_M (resp. J_L , J_R) satisfies $B \leq \tilde{B}^{O(k)}$, where \tilde{B} is an upper bound on the Schmidt rank of H_M (resp. H_L , H_R) at every cut.

⁶ η_0 and η_1 may be chosen as the ground state energy and first excited energy of H respectively, but they need not.

5. The Schmidt rank of K with respect to the cuts (u_1, u_1+1) and (u_2, u_2+1) satisfies $D \leq (dk)^{O(\ell+k/\ell)}$.

Proof. Item 1. follows from the definition of $K = T_k(H)$ as a polynomial in H (see Definition 6.3). For item 2. and item 3 we use the following properties of T_k (see e.g. [AKLV13] and [KAAV15, Lemma B.1] for a proof):

$$|T_k(x)| \leq 1 \quad \text{for } |x| \leq 1, \quad (5)$$

$$|T_k(x)| \geq \frac{1}{2} \exp \left(2k \sqrt{\frac{|x|-1}{|x|+1}} \right) \quad \text{for } |x| \geq 1, \quad (6)$$

$$T_k(x) = \frac{1}{2} (x + \sqrt{x^2 - 1})^k + \frac{1}{2} (x - \sqrt{x^2 - 1})^k \quad \text{for } |x| \geq 1. \quad (7)$$

The fact that eigenvectors with eigenvalue η_0 are mapped to fixed points of K follows from $P_k(\eta_0) = 1$. Next suppose $|\psi\rangle$ is an eigenvector of H with eigenvalue $\eta_0 + \delta$ where $|\delta| < \eta_1 - \eta_0$. From (7) we see $|T_k(x + \delta) - T_k(x)| = O(k\delta / \min(x^2 - 1, x \pm \sqrt{x^2 - 1}))$ as long as $x, x + \delta \leq -1$. Taking into account the scaling used to define P_k ,

$$|P_k(\eta_0 + \delta) - P_k(\eta_0)| = O\left(\frac{1}{\tilde{P}_k(\eta_0)} \frac{k\delta}{\gamma \|H\|}\right) = O\left(\frac{\delta k}{\gamma \|H\|}\right) e^{-2k \sqrt{\frac{\gamma}{\|H\| - \eta_0}}},$$

where the last inequality uses (6). Item 3 follows by combining (5) and (6).

Item 4. is immediate since K is computed as a linear combination of j -th powers of H for $j \leq k$.

Finally, item 5 follows from Proposition 6.5. \square

Theorem 6.4 provides us with a powerful recipe for constructing good AGSP. To minimize the Schmidt rank at a cut $(u, u + 1)$ for $u \in \{u_1, u_2\}$ we should take $k = \Theta(\ell^2)$, which gives a bound of $(dk)^{O(\sqrt{k})}$, a much better bound than the naive $d^{O(k)}$. To guarantee a small Δ we should choose k large enough to ensure that $e^{-4k \sqrt{\gamma/\|H\|}}$ remains small, which requires the Hamiltonian to have a small norm. This is the role of the different truncation schemes that are presented in the following subsections.

For algorithmic purposes it is important that the Chebyshev AGSP can be constructed efficiently once one is given MPO representations for the truncated part of the Hamiltonian. The following proposition states that this is possible.

Proposition 6.5. *Let H be a Hamiltonian having a decomposition as in (4), and K the associated degree- k Chebyshev AGSP. Assume that H_M (but not necessarily H_L or H_R) is specified by an MPO with bond dimensions at most \tilde{B} each.*

Then there exists $D \leq (dk)^{O(\ell+k/\ell)}$ such that a family of D^2 MPO $\{A_1, \dots, A_{D^2}\}$ of bond dimension at most \tilde{B}^k each such that there exists B_1, \dots, B_{D^2} with $K = \sum A_i \otimes B_i$ can be computed in time $nD^2 \tilde{B}^{O(k)}$. Here the A_i act on qudits $\{u_1, u_1 + 1, \dots, u_2\}$ and the B_i on the remaining qudits. This computation does not require knowledge of η_0, η_1 .

Furthermore, if H_L and H_R are also given as MPO with bond dimension at most \tilde{B} then the B_i can be computed as well.

Proof. The proof follows from a close examination of the proof of [AKLV13, Lemma 4.2], which establishes a claim analogous to item 5 in Theorem 6.4.⁷ Adapting to our setting (where there are two cuts

⁷To follow the ensuing argument it may be helpful to translate the notation used for the indices in [AKLV13, Lemma 4.2] to the notation used here as follows: $s \rightarrow 2\ell - 2$, $\ell \rightarrow k$, $k \rightarrow j$.

to consider simultaneously) the argument made in [AKLV13] shows that in order to obtain an MPO for K it suffices to include in the set $\{A_1, \dots, A_{D^2}\}$ MPO representations for operators $P_{i_1 i_2, k j_1 j_2}(Z)$ where $i_1 \in \{u_1 - \ell, \dots, u_1 + \ell - 1\}$, $i_2 \in \{u_2 - \ell, \dots, u_2 + \ell - 1\}$, $j_1, j_2 \in \{0, \dots, k + 2\ell - 2\}$ and Z is an $(4\ell - 4)$ -tuple of complex variables which takes on $\binom{k-j_1+2\ell-2}{2\ell-2} \binom{k-j_2+2\ell-2}{2\ell-2}$ possible values. For our purposes, a random choice of such values, e.g. distributed uniformly on the unit circle, will lead to a correct construction with probability 1 (i.e. only depending on the number of digits of accuracy). We argue below that for each $P_{i_1 i_2, k j_1 j_2}(Z)$ one can efficiently construct an explicit set of MPO $\{A_\alpha\}$, where $1 \leq i \leq \binom{k+j_1+1}{2j_1+1} \binom{k+j_2+1}{2j_2+1} d^{2(j_1+j_2)+4\ell}$, such that there exists B_i for which $\sum A_i \otimes B_i$ is an MPO for $P_{i_1 i_2, k j_1 j_2}(Z)$. This will lead to the claimed bounds as

$$\sum_{i_1=u_1-\ell}^{u_1+\ell-1} \sum_{i_2=u_2-\ell}^{u_2+\ell-1} \sum_{j_1, j_2=0}^{\lfloor k/2\ell \rfloor} d^{2(j_1+j_2)+4\ell} \cdot \binom{k-j_1+2\ell-2}{2\ell-2} \binom{k-j_2+2\ell-2}{2\ell-2} \cdot \binom{k+j_1+1}{2j_1+1} \binom{k+j_2+1}{2j_2+1}$$

can be crudely bounded by $(dk)^{O(\ell+k/\ell)}$.

Fix i_1, i_2 and recall that $P_{i_1 i_2, k j_1 j_2}(Z)$ is defined as the sum of those terms in the expansion of $(H_L + \dots + H_{i_1} + \dots + H_{i_2} + \dots + H_R)^k$ which contain exactly j_1 (resp. j_2) occurrences of H_{i_1} (resp. H_{i_2}). There are $\binom{k+j_1+1}{2j_1+1} \binom{k+j_2+1}{2j_2+1}$ such terms. By cutting to the left of i_1 and right of i_2 we can efficiently construct at most $d^{2(j_1+j_2)}$ MPO which, properly combined, would give an MPO for the corresponding product. Finally we cut these MPO further so as to make the separation be to the left of u_1 and right of u_2 (or complete them appropriately, depending on whether $i_1 \leq u_1$ or $i_1 > u_1$, and similarly for i_2 with respect to u_2). This last step multiplies the number of MPO by at most $d^{4\ell}$ (where we use $|u_1 - i_1|, |u_2 - i_2| \leq \ell$), giving the claimed bound. \square

6.3 The frustration-free case

Our first truncation scheme applies to the case of a local Hamiltonian $H = \sum_i h_i$ that is frustration-free with a spectral gap $\gamma > 0$ and a unique ground state $|\Gamma\rangle$. Replacing each h_i by the projection on its range preserves the ground state and, given our normalization assumption $0 \leq h_i \leq \mathbb{1}$, can only increase the spectral gap; thus we may without loss of generality assume that each h_i is a projection.

We define a truncated version of H based on the detectability lemma from [AAVL11] as follows.

Definition 6.6 (Truncated Hamiltonian in the frustration-free case). *Suppose given a local Hamiltonian H such that $H = H_J + H_{\bar{J}}$ where $H_J = h_{j_0} + h_{j_0+1} + \dots + h_{j_1-1}$ is a local Hamiltonian acting on a contiguous set of qudits $J = \{j_0, j_0 + 1, \dots, j_1\}$. Let J_e (resp. J_o) denote the subset of indices $i \in J$ that are even (resp. odd). Define $H_{J,e} := \sum_{i \in J_e} h_i$ and $H_{J,o} := \sum_{i \in J_o} h_i$. Then the truncation of H_J is given by $\tilde{H}_J := \tilde{H}_{J,e} + \tilde{H}_{J,o}$, where*

$$\tilde{H}_{J,e} := \mathbb{1} - \otimes_{i \in J_e} (\mathbb{1} - h_i), \quad \tilde{H}_{J,o} := \mathbb{1} - \otimes_{i \in J_o} (\mathbb{1} - h_i). \quad (8)$$

The truncated Hamiltonian \tilde{H} associated to region J is given by

$$\tilde{H} := \tilde{H}_J + H_{\bar{J}}. \quad (9)$$

Clearly, $\tilde{H}_{J,e}$ and $\tilde{H}_{J,o}$ are projectors and hence their norm is 1. In addition, they are the sum of the identity operator and a product of non-overlapping local terms, and as such, their Schmidt rank is at most $d^2 + 1$ across any cut. We show that \tilde{H} has the same ground state as H , as well as a large spectral gap. This is done through the detectability lemma and a new converse to the lemma stated below.

Definition 6.7 (The detectability lemma operator in 1D). *Let $H = h_1 + \dots + h_{n-1}$ be a 1D nearest-neighbor Hamiltonian such that each h_i is a projector. Then the DL operator of H is defined by*

$$\text{DL}(H) := \otimes_i (\mathbb{1} - h_{2i}) \otimes_i (\mathbb{1} - h_{2i+1}).$$

Note that the operator $\text{DL}(H)$ is in general not Hermitian. The usefulness of the definition comes primarily from the detectability lemma [AAVL11]:

Lemma 6.8 (The detectability lemma). *Let h_1, \dots, h_m be projectors such that each h_i commutes with all but at most g other h_j , and let $H := \sum_i h_i$. For any state $|\psi\rangle$ let $|\phi\rangle := \prod_i (\mathbb{1} - h_i)|\psi\rangle$, where the product is taken in any order. Then*

$$\| |\phi\rangle \|^2 \leq \frac{1}{\varepsilon_\phi / g^2 + 1}, \quad \text{where} \quad \varepsilon_\phi := \frac{1}{\| |\phi\rangle \|^2} \langle \phi | H | \phi \rangle. \quad (10)$$

The version of the detectability lemma stated above is stronger and more general than the one appearing in [AAVL11]. It also has a much simpler proof, which is given in Ref. []. **(Itai: Fill in as soon as we post on arXiv)**

In addition to the detectability lemma, we will use a converse statement which gives a lower bound on the norm of $\text{DL}(H)|\psi\rangle$:

Lemma 6.9 (Converse of detectability lemma). *Let $H = \sum_{i=1}^{n-1} h_i$ be a 1D nearest-neighbor Hamiltonian such that each h_i is a projector. Then for any eigenvector $|\psi\rangle$ of H ,*

$$\| \text{DL}(H)|\psi\rangle \|^2 \geq 1 - 4\varepsilon'_\psi, \quad \text{where} \quad \varepsilon'_\psi := \langle \psi | H | \psi \rangle. \quad (11)$$

Also the proof of this lemma is given in Ref. [].

With these two lemmas at hand we show the following.

Theorem 6.10. *The truncated Hamiltonian \tilde{H} from Definition 6.6 satisfies the following:*

1. \tilde{H} is frustration free and has the same ground state $|\Gamma\rangle$ as H .
2. The Schmidt rank of \tilde{H} at every cut is at most $d^2 + 2$.
3. \tilde{H} has a spectral gap $\tilde{\gamma} = \Omega(\gamma)$.

Proof. Property 1. follows from the definition. For property 2. note first that the Schmidt rank of every operator on two d -dimensional qudits is at most d^2 . This implies that the Schmidt rank of \tilde{H} at every cut in \bar{J} is at most $d^2 + 2$: we get a d^2 contribution from the local term that is defined on the cut and the extra 2 comes from terms to the right/left of the cut. Consider now a cut between $i, i + 1$ for an even i that is in J . Since i is even $\tilde{H}_{J,e}$ will contribute at most d^2 , and $\tilde{H}_{J,o}$ at most 1. The terms in $H_{\bar{J}}$ contribute at most 1 as well, giving the claimed bound of $d^2 + 2$.

To prove 3. let $|\psi\rangle$ be orthogonal to $|\Gamma\rangle$. By the detectability lemma applied to H , $\| \text{DL}(H)|\psi\rangle \| \leq \frac{1}{\gamma/4+1}$. By the converse of the detectability lemma applied to \tilde{H} , $\| \text{DL}(\tilde{H})|\psi\rangle \| \geq 1 - 4\tilde{\gamma}$. Since by construction $\text{DL}(H) = \text{DL}(\tilde{H})$, this implies

$$\tilde{\gamma} \geq \frac{1}{4} \left(1 - \frac{1}{\gamma/4+1} \right),$$

from which the claim follows. □

We end this section by giving the proof of Theorem 4.3 from Section 4.1.

Proof of Theorem 4.3. Let $J_L = \{1, \dots, i_1\}$, $J_M = \{i_1 + 1, \dots, i_2\}$ and $J_R = \{i_2 + 1, \dots, n\}$ be the set of qudits contained in \mathcal{H}_L , \mathcal{H}_M and \mathcal{H}_R respectively. We first construct a suitable AGSP K , from which the operators A_i will be derived.

The first step in the construction of K consists in truncating the Hamiltonian associated to each of the three regions. For this, introduce a width parameter

$$\ell = \tilde{\Theta}(\gamma^{-1} \log^2 d), \quad (12)$$

and define a Hamiltonian \tilde{H} by applying the truncation scheme described in Definition 6.6 thrice, to the regions $J_L = \{1, \dots, i_1 - \ell - 1\}$, $J_M = \{i_1 + \ell + 1, \dots, i_2 - \ell - 1\}$ and $J_R = \{i_2 + \ell + 1, \dots, n\}$ respectively (provided each region is non-empty). Based on Theorem 6.10 the resulting truncated Hamiltonian \tilde{H} has norm $O(1)$, the same ground state as H , and a spectral gap $\tilde{\gamma} = \Theta(\gamma)$.

K is obtained by applying Definition 6.3 to \tilde{H} with $\eta_0 = 0$, $\eta_1 = \tilde{\gamma}$ and $k = \Theta(\ell^2)$. The bound on D follows from Theorem 6.4, using which one can also verify that the desired trade-off $D^{12} \Delta < \frac{1}{2000}$ will be achieved provided the right choice of constants is made in the choice of ℓ .

By Theorem 6.10 \tilde{H} can be represented as an MPO with bond dimension $O(d^2)$, from which it follows that we can compute a decomposition $K = \sum L_i \otimes A_i \otimes R_i$ where each A_i has bond dimension $O(d^k) = e^{\tilde{\Theta}(\gamma^{-2} \log^5 d)}$.

The claim on the running time follows from the estimates provided in Proposition 6.5. \square

6.4 Hard truncation

In this section we introduce a scheme of *hard truncation* that is appropriate (though not efficient) for truncating the norm of an arbitrary local Hamiltonian in a certain region J , while preserving its low-energy eigenspace $H_{\varepsilon_0, \varepsilon_0 + \eta}$. The basic idea is to replace $H \mapsto H\Pi_{\leq \varepsilon_0 + t} + (\varepsilon_0 + t)\Pi_{> \varepsilon_0 + t}$, where $\Pi_{\leq t}$ projects onto the span of eigenvectors of H with eigenvalue less than t , $\Pi_{> \varepsilon_0 + t} := \mathbb{1} - \Pi_{\leq \varepsilon_0 + t}$, and t is chosen to be large enough with respect to η .

Definition 6.11 (Hard truncation). *Let $t > 0$, $H = H_J + H_{\bar{J}}$ where $H_J = h_{j_0} + h_{j_0+1} + \dots + h_{j_1-1}$ is a local Hamiltonian acting on a contiguous set of qudits $J = \{j_0, j_0 + 1, \dots, j_1\}$, and let ε_J be the ground energy of H_J . Let Π_- be the projector onto the span of all eigenvectors of H_J with eigenvalue less than $\varepsilon_J + t$, and $\Pi_+ := \mathbb{1} - \Pi_-$. Then the hard truncation of H_J is given by*

$$\tilde{H}_J := H_J \Pi_- + (t + \varepsilon_J) \Pi_+ \quad (13)$$

and the hard-truncated Hamiltonian \tilde{H}_t associated to the region J is

$$\tilde{H}_t = \tilde{H}_J + H_{\bar{J}}.$$

We now show that truncating a n -qubit Hamiltonian on a subset J of the qubits leads to a truncated Hamiltonian whose low-energy space is close to that of the original Hamiltonian. The main tool in proving this result is Theorem 2.6 of Ref. [AKL14], a generalization and strengthening of the truncation result that appeared in Ref. [AKLV13]. Adapting it to the current setting it can be formulated as follows.

Proposition 6.12 (Adapted from Theorem 2.6 of Ref. [AKL14]). *For any $\eta > 0$ let $\Pi_{\leq \eta}$ denote the projector on the span of all eigenvectors of H with eigenvalue at most η , and similarly $\tilde{\Pi}_{\leq \eta}$ for \tilde{H}_t . Let $\varepsilon_0 \leq \varepsilon_1 \leq$*

$\varepsilon_2 \leq \dots$ and $\tilde{\varepsilon}_0 \leq \tilde{\varepsilon}_1 \leq \tilde{\varepsilon}_2 \dots$ be the sorted eigenvalues of H and \tilde{H}_t , respectively where eigenvalues appear with multiplicity. For any $\eta > 0$, let

$$\xi = e^{-C_1(t-\eta)+C_0}, \quad (14)$$

with $C_0 = 24$ and $C_1 = \frac{1}{8}$. Then the following holds:

1. $\|(H - \tilde{H}_t)\Pi_{\leq \varepsilon_0 + \eta}\| \leq \xi$ and $\|(H - \tilde{H}_t)\tilde{\Pi}_{\leq \varepsilon_0 + \eta}\| \leq \xi$,
2. For all j for which $\varepsilon_j \leq \varepsilon_0 + \eta$, we have $\varepsilon_j - \xi \leq \tilde{\varepsilon}_j \leq \varepsilon_j$.

Proof. The proposition follows from Theorem 2.6 in Ref. [AKL14] by using $\lambda = \frac{1}{8}$ and the fact that $\varepsilon_0 \leq \tilde{\varepsilon}_0 + 2$ to bound $\Delta\tilde{\varepsilon}$ by $\Delta\varepsilon + 2$. Here we can take $|\partial L| = 2$ since there are 2 boundary terms connecting the truncated region J and the rest of the system. \square

The following lemma summarizes the approximation properties of the hard truncation procedure that will be important for us.

Lemma 6.13. For any $\eta > 0$, let $T_\eta = H_{[\varepsilon_0, \varepsilon_0 + \eta]}$ be the low-energy eigenspace of H , $J = \{j_0, \dots, j_1\}$ a contiguous subset of qudits and \tilde{H}_t the associated hard-truncated Hamiltonian, with corresponding low-energy eigenspace $\tilde{T}_\eta = \tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \eta]}$. Let ξ be as defined in (14). Then the following hold for any $t > \eta$:

1. The ground energy $\tilde{\varepsilon}_0$ of \tilde{H}_t satisfies $\varepsilon_0 - Ce^{-c(t-\eta)} \leq \tilde{\varepsilon}_0 \leq \varepsilon_0$ for some universal constants C, c .
2. For any $\delta > 0$ there is

$$\eta' = \eta + \sqrt{\frac{\eta}{\delta}} e^{-\Omega(t-\eta)}$$

such that the subspace $\tilde{T}_{\eta'}$ is δ -close to T_η , and $T_{\eta'}$ is δ -close to \tilde{T}_η .

Proof. The first item follows directly from the second item in Proposition 6.12. For the second item, we prove that $\tilde{T}_{\eta'}$ is δ -close to T_η , the proof of the second relation being identical. Fix a small width parameter h (to be specified later) and let $|\psi\rangle = \sum_i \beta_i |\psi_i\rangle$ be supported on eigenvectors $|\psi_i\rangle$ of H with eigenvalue $\mu_i \in [\lambda - h, \lambda + h]$ with $\lambda \leq \varepsilon_0 + \eta$. Then $\|H|\psi\rangle - \lambda|\psi\rangle\| \leq h$. Decompose $|\psi\rangle = \sum \alpha_i |\phi_i\rangle$, where for each i , $|\phi_i\rangle$ is an eigenvector of \tilde{H}_t with associated eigenvalue $\tilde{\lambda}_i$. Using the first item in Proposition 6.12,

$$\begin{aligned} \sum_i |\alpha_i|^2 |\lambda - \tilde{\lambda}_i|^2 &\leq (\|(H - \tilde{H})|\psi\rangle\| + \|(H - \lambda\mathbb{1})|\psi\rangle\|)^2 \\ &\leq (e^{-\Omega(t-\eta)} + h)^2. \end{aligned}$$

By Markov's inequality it follows that for any $\delta > 0$

$$\|\tilde{\Pi}_{>\lambda+\delta}|\psi\rangle\| \leq \frac{e^{-\Omega(t-\eta)} + h}{\delta}.$$

Any $|\psi\rangle$ in T_η can be written as a linear combination $|\psi\rangle = \sum_j \beta_j |h_j\rangle$ with each $|h_j\rangle$ supported on eigenvectors of H with eigenvalue in a small window of width $2h$, and the number of terms at most $\lceil \frac{\eta - \varepsilon_0}{2h} \rceil$. Thus

$$\begin{aligned} \|\tilde{\Pi}_{>\varepsilon_0 + \eta'}|\psi\rangle\| &\leq \sum_j |\beta_j| \|\tilde{\Pi}_{>\varepsilon_0 + \eta'}|h_j\rangle\| \\ &\leq \sqrt{\frac{\eta}{2h}} \frac{e^{-\Omega(t-\eta)} + h}{\eta' - \eta}. \end{aligned}$$

Choosing $h = e^{-\Theta(t)}$, we see that the choice of η' made in the statement of the lemma suffices to ensure that this quantity is at most $\sqrt{\delta}$, as desired. \square

The combination of Theorem 6.4, Proposition 6.12, and Lemma 6.13 yield a construction that starts with a local Hamiltonian H , produces a truncated Hamiltonian \tilde{H} with low energy space close to that of H along with a spectral AGSP K for \tilde{H} with a good trade-off between the parameters D and Δ .

Corollary 6.14. *Let H be a 1D local Hamiltonian with ground energy ε_0 , and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the n -qudit space in contiguous regions. For any integer ℓ and $t > 0$ there exists a Hamiltonian \tilde{H} such that for any $\varepsilon_0 < \eta_0 < \eta_1$ there is a spectral AGSP K for $(\tilde{H}, \eta_0, \eta_1)$ with the following properties.*

1. **Good parameters:** K has the parameters $D = (d\ell)^{O(\ell)}$ and $\Delta = e^{-\Omega(\frac{\ell^2}{\sqrt{t+\ell}}\sqrt{\eta_1-\eta_0})}$,

2. **Closeness:**

(a) *There are universal constants $C, c > 0$ such that*

$$0 \leq \varepsilon_i - \tilde{\varepsilon}_i \leq Ce^{-c(t-(\varepsilon_i-\varepsilon_0))} \quad (15)$$

where $\varepsilon_i, \tilde{\varepsilon}_i$ are the i -th smallest (counted with multiplicity) eigenvalues of H, \tilde{H} respectively.

(b) *The space $H_{[\varepsilon_0, \eta_1]}$ is δ -close to $\tilde{H}_{[\tilde{\varepsilon}_0, \eta_0]}$ and $\tilde{H}_{[\tilde{\varepsilon}_0, \eta_1]}$ is δ -close to $H_{[\varepsilon_0, \eta_0]}$, for*

$$\delta = \Theta\left(\frac{\eta_0 - \varepsilon_0}{(\eta_1 - \eta_0)^2}\right)e^{-\Omega(t-(\eta_0-\varepsilon_0))}. \quad (16)$$

Proof. Let $J_L = \{1, \dots, i_1\}$, $J_M = \{i_1 + 1, \dots, i_2\}$ and $J_R = \{i_2 + 1, \dots, n\}$ be the set of qudits contained in $\mathcal{H}_L, \mathcal{H}_M$ and \mathcal{H}_R respectively. We define the truncated Hamiltonian \tilde{H} by applying the hard truncation transformation described in Definition 6.11 thrice, to the regions $J_L = \{1, \dots, i_1 - \ell - 1\}$, $J_M = \{i_1 + \ell + 1, \dots, i_2 - \ell - 1\}$ and $J_R = \{i_2 + \ell + 1, \dots, n\}$ respectively (provided each region is non-empty). The resulting truncated Hamiltonian $\tilde{H} = \tilde{H}_t$ has norm $O(\ell + t)$.

Applying Lemma 6.13 thrice in sequence, for the three truncations performed, it follows that the sorted eigenvalues of \tilde{H} satisfy (15). Eq. (16) similarly follows from item 2. in Lemma 6.13

Finally we define the AGSP K by applying the Chebyshev polynomial construction from Definition 6.3 to \tilde{H} with a choice of $k = \ell^2$. The bounds on Δ and D follow directly from item 3. and 5. from Theorem 6.4 respectively. \square

6.4.1 AGSP constructions

In this section we provide details for the constructions of AGSP whose existence is claimed in Theorem 3.8 and Theorem 5.1 respectively. Both constructions are derived from Corollary 6.14, merely requiring different choices of parameters.

Proof of Theorem 3.8. Let $\eta_0 = \varepsilon_0 + \gamma/10$ and $\eta_1 = \varepsilon_0 + 9\gamma/10$. Provided the implied constants are chosen large enough, setting $t = \tilde{O}(\ell)$ and $t > \tilde{O}(\frac{1}{\gamma} \log^2(d/\gamma))$ in Corollary 6.14 gives $D^{12}\Delta < \frac{1}{2000}$. Due to the gap assumption it holds that $T = H_{[\varepsilon_0, \eta_0]} = H_{[\varepsilon_0, \eta_1]}$. The choice of t above also insures that the right-hand side of (15) is smaller than $\frac{1}{10}\gamma$ and the right hand side of (16) is smaller than .005, in which case \tilde{H} has

a spectral gap between η_0 and η_1 , so that $\tilde{H}_{[\tilde{\varepsilon}_0, \eta_0]} = \tilde{H}_{[\tilde{\varepsilon}_0, \eta_1]}$. Then item 2(b) in the corollary implies that $\tilde{H}_{[\tilde{\varepsilon}_0, \eta_0]}$ and T are mutually .005-close.

The operators $\{A_i\}$ claimed in Theorem 3.8 are defined as the $\{M_i\}$ in a decomposition $K = \sum_{i=1}^{D^2} L_i \otimes M_i \otimes R_i$ associated to the factorization $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ of the AGSP from Corollary 6.14. Lemma 6.2 gives the desired quantitative tradeoff between the increase in dimension of a viable set and its increase in overlap, when acted upon by the $\{M_i\} = \{A_i\}$. Thus we obtain Theorem 3.8 with $\tilde{T} = \tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \frac{1}{10}\gamma]}$. \square

Proof of Theorem 5.1. The main difference with the proof of Theorem 3.8 is that the parameter corresponding to the gap γ is replaced by the quantity $\frac{\mu}{\log n}$. The proof of the first two items claimed in the theorem then closely mirrors that of Theorem 3.8.

It only remains to verify the third item. Despite having the desired AGSP, unlike in the gapped case we cannot hope to improve the quality of the viable set S for all of $\tilde{T} = \tilde{H}_{[0, \eta' - \frac{\mu}{3 \log n}]}$ by the application of the AGSP K^k . However, if we view S as a viable set for the smaller $\tilde{T}_- = \tilde{H}_{[0, \eta' - \frac{2\mu}{3 \log n}]} \subseteq \tilde{T}$, we now have an effective AGSP with respect to \tilde{T}_- and the orthogonal complement of the larger \tilde{T} and we can proceed as if in the presence of a small spectral gap of $\frac{\mu}{3 \log n}$. To see this, fix any vector $|\psi\rangle \in \tilde{T}_-$. Lemma 3.5 shows that there exists a $|w\rangle \in S$ such that $|w\rangle = c|\psi\rangle + |\psi^\perp\rangle$ for some $|\psi^\perp\rangle$ orthogonal to \tilde{T} , and $c \geq (1 - \delta)$. This brings us in line with the proof of Lemma 6.2 and we can use the same analysis to show that applying K^k improves the parameter of the viable set S from δ to the desired $\delta' = \frac{\Delta^k}{(1-\delta)^2}$. \square

6.5 Soft truncation

In this section we introduce a different scheme of *soft truncation* that reduces the norm of a local Hamiltonian H in a certain region J in a way that the truncated operator can be well-approximated by an MPO with small bond dimension. The soft truncation procedure approximates (to the first order; see Definition 6.16 for details) $H_{[\varepsilon_0, \varepsilon_0 + t]}$ via the operator $t(1 - e^{-H/t})$, and then leverages the truncated cluster expansion [Has06, KGK⁺14] and its matrix product operator (MPO) representation from [MSVC15, Section IV] to approximate $e^{-H/t}$ by an operator with polynomial Schmidt rank.

In hard truncation (Definition 6.11) the operator $\Pi_{\leq \varepsilon_0 + t} H + (\varepsilon_0 + t) \Pi_{> \varepsilon_0 + t}$ is used. This can be written as $g_t(H)$, where $g_t(x)$ is defined by $g_t(x) := x$ for $x \leq \varepsilon_0 + t$ and $g_t(x) := t$ for $x > \varepsilon_0 + t$. The main idea of soft truncation is to replace this non-smooth function by the infinitely differentiable function

$$h_{t', t}(x) := t \left(f_t(x) + \frac{f_t(x)^2}{2} + \dots + \frac{f_t(x)^{t'}}{t'} \right), \quad \text{where} \quad f_t(x) := 1 - e^{-x/t}, \quad (17)$$

which results in an operator $h_{t', t}(H)$ that can be given an efficient representation as an MPO. The following are basic properties of $h_{t', t}$.

Lemma 6.15. *For any integers $t', t \geq 1$ and $x \geq 0$,*

$$|h_{t', t}(x) - x| \leq \frac{t}{t'} \left(\frac{x}{t} \right)^{t'}, \quad \text{and} \quad |h_{t', t}(x)| \leq t \ln(t').$$

Proof. Let $g_t(y) = -t \ln(1 - y)$, so that $g_t(f_t(x)) = x$ for any $x \in [0, \infty)$. The function $h_{t', t}$ contains the first t' terms of the Taylor expansion of g_t around 0, applied to $f_t(x)$, and the first inequality follows from Taylor's theorem and $f_t(x) \leq x$ for all x . The second inequality follows since $f_t(x) \leq 1$ for all x . \square

In addition to the truncation parameters t and t' the soft truncation construction is parametrized by a region J , which specifies the set of local terms on which truncation is to be performed, and an energy ε'_J which is meant to be an approximation to the ground state energy of the restriction H_J of H to J .

Definition 6.16 (Soft truncation). *Let $H = H_J + H_{\bar{J}}$ be a 1D Hamiltonian, where $H_J = h_{j_0} + \dots + h_{j_1-1}$ acts on a contiguous set $J = \{j_0, \dots, j_1\}$ of qudits. Let ε_J be the ground energy of H_J , and ε'_J an approximation to ε_J satisfying $\varepsilon_J - 10 \leq \varepsilon'_J \leq \varepsilon_J$. For given truncation parameters $t \geq t' \geq 1$, the soft truncation of H_J is given by*

$$\tilde{H}_J := \varepsilon'_J \mathbb{1} + h_{t',t}(H_J - \varepsilon'_J \mathbb{1}),$$

and the soft-truncated Hamiltonian H associated to region J is

$$\tilde{H}_{t',t} := \tilde{H}_J + H_{\bar{J}}.$$

The following lemma shows that for sufficiently large t and t' , $\tilde{H}_{t',t}$ provides a good approximation to the lower part of the spectrum of H .

Lemma 6.17. *Let $H = H_J + H_{\bar{J}}$ be a local 1D Hamiltonian. Given truncation parameters $t \geq t' \geq 2$, the soft-truncated Hamiltonian $\tilde{H}_{t',t}$ satisfies $\tilde{H}_{t',t} \leq H$ and for any eigenvector $|\psi\rangle$ of H with energy λ (resp. $|\phi\rangle$ of \tilde{H} with energy $\mu \leq t$) it holds that*

$$\lambda - O\left(\frac{(\lambda - \varepsilon)^{t'}}{t't^{t'-1}}\right) \leq \langle \psi | \tilde{H}_{t',t} | \psi \rangle \leq \lambda \quad \text{and} \quad \mu \leq \langle \phi | H | \phi \rangle \leq \mu + O\left(\frac{(2(\mu - \varepsilon))^{t'}}{t't^{t'-1}}\right), \quad (18)$$

where $\varepsilon = \varepsilon_{\bar{J}} + \varepsilon'_J$. In addition, if H is gapped with gap γ then provided $t = \Omega(\gamma^{-1})$, $\tilde{H}_{t',t}$ is gapped with gap $\gamma/2 \leq \tilde{\gamma} \leq 2\gamma$.

For $\eta > 0$ let $T_\eta = H_{[\varepsilon_0, \varepsilon_0 + \eta]}$ (resp. $\tilde{T}_\eta = \tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \eta]}$) be the span of all eigenvectors of H (resp. $\tilde{H}_{t',t}$) with associated eigenvalue in the indicated range. Then for any $\eta, \delta > 0$ there is

$$\eta' = \eta + O\left(\left(\frac{\eta + 10}{t}\right)^{t'-1} \frac{1}{t'\sqrt{\delta}}\right)$$

such that the subspace $\tilde{T}_{\eta'}$ is δ -close to T_η and $T_{\eta'}$ is δ -close to \tilde{T}_η .

Proof. From Definition 6.16,

$$\tilde{H}_{t',t} - H = h_{t',t}(H_J - \varepsilon'_J \mathbb{1}) - (H_J - \varepsilon'_J \mathbb{1}). \quad (19)$$

Using the first bound from Lemma 6.15, we get that for any vector $|\psi\rangle$,

$$|\langle \psi | \tilde{H}_{t',t} | \psi \rangle - \langle \psi | H | \psi \rangle| \leq \frac{1}{t't^{t'-1}} \langle \psi | (H_J - \varepsilon'_J \mathbb{1})^{t'} | \psi \rangle. \quad (20)$$

Furthermore,

$$\begin{aligned} H_J - \varepsilon'_J \mathbb{1} &\leq H_{\bar{J}} - \varepsilon_{\bar{J}} \mathbb{1} + H_J - \varepsilon'_J \mathbb{1} \\ &= H - (\varepsilon_{\bar{J}} + \varepsilon'_J) \mathbb{1}, \end{aligned}$$

which combined with (20) and $H_J - \varepsilon'_J \mathbb{1} \geq 0$ proves the first two inequalities in (18); the other two are obtained in the same way using in addition $x \leq 2h_{t',t}(x)$ for $0 \leq x \leq t$. The relations between the spectral gaps of H and $\tilde{H}_{t',t}$ follow from these inequalities.

Starting from (19), squaring both sides and using (the square of) the bound from Lemma 6.15 we get the operator inequality

$$(\tilde{H}_{t',t} - H)^2 \leq \frac{1}{(t')^{2t'2^{t'-2}}} (H_J - \varepsilon'_J \mathbb{1})^{2t'}. \quad (21)$$

Let $\bar{H}_J = H_J - h_{j_0-1} - h_{j_1}$, so that \bar{H}_J and H_J commute. Using $\bar{H}_J + (2 - \varepsilon_J) \mathbb{1} \geq 0$,

$$\begin{aligned} (H_J - \varepsilon'_J \mathbb{1})^{2t'} &\leq (H_J - \varepsilon'_J \mathbb{1} + \bar{H}_J + (2 - \varepsilon_J) \mathbb{1})^{2t'} \\ &\leq ((H - \varepsilon) \mathbb{1} + 10 \mathbb{1})^{2t'}. \end{aligned} \quad (22)$$

Let $|\psi\rangle$ be supported on eigenvectors of H with eigenvalues in the range $[\lambda - h, \lambda + h]$ with $\lambda \leq \varepsilon_0 + \eta$ and h a small width parameter. Decompose $|\psi\rangle = \sum_i \alpha_i |\phi_i\rangle$, where for each i , $|\phi_i\rangle$ is an eigenvector of $\tilde{H}_{t',t}$ with associated eigenvalue $\tilde{\lambda}_i$. Thus

$$\begin{aligned} \left(\sum_i |\alpha_i|^2 |\lambda - \tilde{\lambda}_i|^2 \right)^{1/2} &\leq \left\| \sum_i \alpha_i (\lambda - \tilde{\lambda}_i) |\phi_i\rangle \right\| + h \\ &= \|(\tilde{H}_{t',t} - H)|\psi\rangle\| + h \\ &= \langle \psi | (\tilde{H}_{t',t} - H)^2 | \psi \rangle^{1/2} + h \\ &\leq \frac{1}{t't^{t'-1}} \langle \psi | ((H - \varepsilon) \mathbb{1} + 10 \mathbb{1})^{2t'} | \psi \rangle^{1/2} + h \\ &\leq \frac{1}{t't^{t'-1}} (\eta + 10)^{t'} + h, \end{aligned}$$

where the inequality before last follows by combining (21) and (22). Applying Markov's inequality it follows that for any $\delta > 0$

$$\|\tilde{\Pi}_{>\lambda+\delta} |\psi\rangle\| \leq \frac{\frac{1}{t't^{t'-1}} (\eta + 10)^{t'} + h}{\delta}.$$

Any $|\psi\rangle$ in T_η can be written as a linear combination $|\psi\rangle = \sum_j \beta_j |h_j\rangle$ with each $|h_j\rangle$ supported on eigenvectors of H with eigenvalue in a small window of width $2h$, and the number of terms at most $\lceil \frac{\eta - \varepsilon_0}{2h} \rceil$. Thus

$$\begin{aligned} \|\tilde{\Pi}_{>\varepsilon_0+\eta'} |\psi\rangle\| &\leq \sum_j |\beta_j| \|\tilde{\Pi}_{>\varepsilon_0+\eta'} |h_j\rangle\| \\ &\leq \sqrt{\frac{\eta}{2h}} \frac{\frac{1}{t't^{t'-1}} (\eta + 10)^{t'} + h}{\eta' - \eta}. \end{aligned}$$

Choosing $h = \frac{1}{t't^{t'-1}} (\eta + 10)^{t'}$, we see that the choice of η' made in the statement of the lemma suffices to ensure that this quantity is at most δ , as desired. \square

6.5.1 The cluster expansion

In this section we show that the soft-truncated Hamiltonian $\tilde{H}_{t',t}$ (Definition 6.16) can be approximated by an operator with polynomial Schmidt rank, and can be computed efficiently. Our construction is based on the cluster expansion from [Has06, KGK⁺14] in the 1D case, with some small adjustments. We first state the result.

Lemma 6.18. *Let t and $t' < (\ln(2)/2)t$ be truncation parameters and H a n -qudit local Hamiltonian. For any $\xi > 0$ there is an MPO representation \tilde{H}' for the truncated Hamiltonian $\tilde{H} = \tilde{H}_{t,t'}$ such that $\|\tilde{H} - \tilde{H}'\| \leq \xi$ and \tilde{H}' has bond dimension $\text{poly}(t'2^{t'}n/\xi)$ across all bonds. Such an MPO can be constructed in time polynomial in its size.*

Proof. The truncation $h_{t',t}(H)$ can be expressed as a linear combination of $O(t'2^{t'})$ terms of the form $e^{-\beta H}$ for values of β in $\{1/t, \dots, t'/t\}$; moreover the coefficients of the linear combination are at most $O(t'2^{t'})$ each. Using Theorem 6.19 and the assumption $t'/t \leq \ln(2)/2$ each $e^{-\beta H}$ can be approximated, in the operator norm, by an MPO of the form $M_r(H)$ with error less than $\xi/(t'2^{t'})^2$ as long as $r = \Omega(\ln((t')^2 2^{2t'} n^2 / \xi))$. Finally, Theorem 6.20 states that such an MPO with the claimed bond dimension can be found efficiently. \square

Let $H = \sum_{i=1}^{n-1} h_i$ be a 1D, 2-local Hamiltonian on n qudits of dimension d , with $\|h_i\| \leq 1$ (but the h_i are not necessarily non-negative), and let $\beta > 0$ be an inverse temperature. We write the cluster expansion $e^{-\beta H} = \sum_w f(w)$, where w runs over all words on $\{1, \dots, n-1\}$ and $f(w) := \frac{(-\beta)^{|w|}}{|w|!} h_w$ with $h_w := \prod_{i \in w} h_i$. For an integer $r > 0$, let $S_{<r}$ be the set of all those w such that the support of w , the set of qudits on which h_w acts non-trivially, consists of connected components of size smaller than r . Let $M_r(H) := \sum_{w \in S_{<r}} f(w)$ be the “truncated cluster expansion” of $e^{-\beta H}$. The following theorem follows from the proof of Lemma 2 in [KKG⁺14]; we give the proof in Appendix A.3.

Theorem 6.19. *Let β be such that $e^\beta - 1 < 1$. Then the following approximation holds in the operator norm:*

$$\|e^{-\beta H} - M_r(H)\| \leq e^{n^2(e^\beta - 1)^r} - 1.$$

The next theorem states that the operator $M_r(H)$ can be written efficiently as an MPO. This encoding also shows that the operator $M_r(H)$ has a low Schmidt rank. The proof, which is given in Appendix A.3, follows very closely the ideas of [MSVC15, Section IV].

Theorem 6.20. *The r^{th} order cluster expansion $M_r(H)$ of the operator $e^{-\beta H}$ can be written as an MPO of bond dimension $\leq r^2 d^r$ which can be computed in time $nd^{O(r)}$.*

6.5.2 Efficient AGSP constructions

In this section we use the soft truncation scheme and combine it with the Chebyshev polynomial AGSP to derive the efficient spectral AGSP needed for our algorithms.

Proof of Theorem 4.1. Let $J_L = \{1, \dots, i_1\}$, $J_M = \{i_1 + 1, \dots, i_2\}$ and $J_R = \{i_2 + 1, \dots, n\}$ be the set of qudits contained in \mathcal{H}_L , \mathcal{H}_M and \mathcal{H}_R respectively. We construct an AGSP K from which the operators $\{A_i\}$ claimed in the theorem will be derived.

The first step in the construction of K consists in truncating the Hamiltonian associated to each of the three regions. For this, introduce truncation parameters

$$t = \Theta(\ell), \quad t' = 4, \tag{23}$$

a width parameter

$$\ell = \Theta(\gamma^{-1} \log \gamma^{-1}),$$

and define a Hamiltonian $\tilde{H} = \tilde{H}_{t',t}$ by applying the soft truncation transformation described in Definition 6.16 thrice, to the regions $J_L = \{1, \dots, i_1 - \ell - 1\}$, $J_M = \{i_1 + \ell + 1, \dots, i_2 - \ell - 1\}$ and $J_R = \{i_2 + \ell + 1, \dots, n\}$ respectively (provided each region is non-empty). The resulting truncated Hamiltonian \tilde{H} has norm $O(\ell + t \log t') = O(\ell)$. Note that the computation of the complete Hamiltonian \tilde{H} requires estimates for the ground energies of the restriction of H to each of the three regions that are being truncated. We will only need to efficiently compute an MPO for \tilde{H}_M , for which a rough estimate for the ground state energy of H_M , as provided as input to **Generate**, will be sufficient.

The second step is to apply the Chebyshev polynomial from Definition 6.3 to \tilde{H} to obtain the AGSP K . For this we make a choice of degree

$$k = \ell^2 \quad (24)$$

and set the energy parameters η_0 and η_1 to

$$\eta_0 = \varepsilon_0 + \gamma/10, \quad \eta_1 = \varepsilon_0 + 9\gamma/10.$$

We first verify that K as defined is a spectral AGSP with the required properties, and then we show how it can be computed efficiently. By item 2. from Theorem 6.4 the scaling parameter Δ is given by

$$\Delta := 4e^{-4k\sqrt{\frac{8\gamma}{10(\|\tilde{H}\| - (\varepsilon_0 + \gamma/10))}}} = e^{-\Omega(k\sqrt{\frac{\gamma}{(\ell+t)}})}. \quad (25)$$

Furthermore, applying Theorem 6.4 twice, once for the region centered at i_1 and once for the region centered at i_2 , the bond parameter D of K across each of the cuts $(i_1 : i_1 + 1)$ and $(i_2 : i_2 + 1)$ is bounded by

$$D \leq (dk)^{O(\ell+k/\ell)}. \quad (26)$$

Thus

$$D^{12}\Delta = e^{\gamma^{-1}\tilde{O}(\log(\gamma^{-1}))}e^{-\Omega(\gamma^{-1}\log^{3/2}(\gamma^{-1}))}$$

can be made smaller than $\frac{1}{2000}$, as desired, by choosing the implicit constants to be small enough.

Next we apply Lemma 6.17 to evaluate the closeness between the low-energy subspaces of H and \tilde{H} . Since H has a spectral gap the subspace $T_{\gamma/20} = H_{[\varepsilon_0, \varepsilon_0 + \gamma/20]}$ is the ground space T of H , so that setting $\delta = 0.05$ the lemma implies that $\tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \gamma/10]}$ is δ -close to T as long as the constant implied in the definition (23) of the truncation parameter t is large enough. Conversely, we can write $T = T_{\gamma/2} = H_{[\varepsilon_0, \varepsilon_0 + 9\gamma/10]}$, in which case the lemma implies that T is δ -close to $\tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \gamma/10]}$. Thus the two spaces are δ -close, as desired. Furthermore, the claim on the ground state energies of H_M and \tilde{H}_M follows directly from Lemma 6.17 and our choice of t .

Finally we turn to efficiency, and verify that in time $n^{O(k)} = n^{\tilde{O}(\gamma^{-2})}$ one can construct a set of at most D^2 MPO A_1, \dots, A_{D^2} acting on \mathcal{H}_M such that there exists B_1, \dots, B_{D^2} acting on $\mathcal{H}_L \otimes \mathcal{H}_R$ such that the AGSP K can be represented as $K = \sum A_\alpha \otimes B_\alpha$. For this we first need to construct MPO representations for the truncated terms in the Hamiltonian. According to Proposition 6.5 this will follow from a bound \tilde{B} on an MPO representation (and computation time) for the truncated term $\tilde{H}_M = h_{t',t}(H_M)$ associated with the region J_M . This is provided by Lemma 6.18 (applied to $H_M - \varepsilon'_M \mathbb{1}$), which given our choice of parameters t, t' guarantees that an MPO providing inverse polynomial approximation (in the operator norm) to \tilde{H}_M can be efficiently computed that has polynomial bond dimension across all cuts. \square

Proof of Theorem 5.5. The proof is similar to Theorem 4.1, and the construction of \tilde{H} and K are the same except for a different choice of parameters. Here we choose

$$\ell = \Theta\left(\log \frac{n}{\mu} \log \log \frac{n}{\mu}\right), \quad k = \ell^2 \quad \text{and} \quad t = \Theta(\ell), \quad t' = 4. \quad (27)$$

The truncated Hamiltonian $\tilde{H} = \tilde{H}_{t,t'}$ is defined as in the proof of Corollary , and the AGSP K is obtained by applying the Chebyshev polynomial from Definition 6.3 to \tilde{H} , energy parameters η'_0 and η'_1 to

$$\eta'_0 = \tilde{\varepsilon}_0 + \eta_1 - \frac{\mu}{2 \log n}, \quad \eta'_1 = \tilde{\varepsilon}_0 + \eta_1$$

respectively. As a result the parameters D and Δ satisfy

$$D^{12} \Delta = e^{\log \frac{n}{\mu} \tilde{O}(\log \log \frac{n}{\mu})} e^{-\Omega(\log \frac{n}{\mu} \log^{1.5}(\log \frac{n}{\mu}))} = o(1),$$

which can be made less than $\frac{1}{2000}$ by a proper choice of implied constants. The conditions on closeness of T , T_- and \tilde{T} , \tilde{T}_- follow from an application of Lemma 6.17, observing that our choice of truncation parameters t, t' is sufficient to conclude closeness of the appropriate subspaces. The claim on the ground state energies of H_M and \tilde{H}_M follows directly from Lemma 6.17 as well.

Finally, applying Proposition 6.5 and Lemma 6.18 we see that an MPO for the part of K acting on region M can be computed in time $n^{O(k)} = e^{\tilde{O}(\log^3 n)}$. \square

7 Shortcut: bootstrapping proof of the area law for degenerate Hamiltonians

Here we present a guide to a quick proof of the area Law for the case of a Hamiltonian satisfying Assumption (DG), i.e. having a degenerate gapped ground space. For simplicity consider a cut in the middle of the n -qudit chain on which H acts, between the $(n/2)$ -th and $(n/2+1)$ -th qudits. Consider a subspace $W \subseteq (\mathbb{C}^d)^{\otimes n/2}$ of *minimal* dimension q that is supported on the first $n/2$ qudits and is .015-viable for the ground space T . The key step in establishing an area law is to bound the size of q :

Theorem 7.1. *Let H be a local Hamiltonian satisfying assumption (DG). Then there exists a subspace $W \subseteq (\mathbb{C}^d)^{\otimes n/2}$ of dimension $q = O(re^{\tilde{O}(\frac{1}{\gamma} \log^3 d)})$ that is .015-viable for T .*

The proof combines two ingredients. The first is that we can find subspaces of W whose viability scales proportionally with the size of the subspace:

Lemma 7.2. *Given $W \subseteq (\mathbb{C}^d)^{\otimes n/2}$ a .04-viable set for T of size q , for any $s = \Omega(\log q + r \log \frac{q}{s})$ there exists a $W' \subseteq W$ of size s such that W' is $(1 - .8 \frac{s}{q})$ -viable for T .*

Proof. We show that with positive probability a random subspace $W' \subseteq W$ of dimension s is $(1 - \alpha)$ -viable for T with $\alpha = (.8) \frac{s}{q}$. We accomplish this by establishing that with high probability $\|P_{W'}|v\rangle\|^2 \geq \alpha$ for all $|v\rangle \in T$ of unit norm. The proof is exactly the same as the analysis of the random sampling step in the proof of Theorem 3.9, and we omit the details. \square

The second ingredient is the existence of a suitable AGSP. For this use a small alteration of the AGSP described in Theorem 3.8, along with the error reduction Lemma 6.2.

Lemma 7.3. *Let H be a local Hamiltonian satisfying Assumption (DG), and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R$ a decomposition of the n -qudit space in two contiguous blocks. There exists a collection of D^2 operators $\{A_i\}_{i=1}^{D^2}$ acting on \mathcal{H}_M along with a subspace $\tilde{T} \subseteq \mathcal{H}$ such that:*

- $\angle_m(T, \tilde{T}) \leq .005$,

- $D = e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$,
- There is $\Delta > 0$ such that $D^2 \Delta \leq \frac{1}{1400}$ and whenever $S \subseteq \mathcal{H}_M$ is δ -viable for \tilde{T} then $S' = \text{Span}\{\cup_i A_i S\}$ is δ' -viable for \tilde{T} , with $\delta' = \frac{\Delta}{(1-\delta)^2}$.

Proof. There are two differences between this AGSP and the one guaranteed in Theorem 3.8. The first is that we are concerned with the decomposition across one cut instead of two, which can be analyzed by thinking of \mathcal{H}_L in Theorem 3.8 as being empty and \mathcal{H}_M in the theorem as \mathcal{H}_L here. The second is the condition $D^2 \Delta < \frac{1}{1400}$, which is implied by $D^{12} \Delta < \frac{1}{2000}$. \square

Proof of Theorem 7.1. Starting with the W of size q , the first condition in Lemma 7.3 together with Lemma 3.7 establishes that W is .04-viable for \tilde{T} . Let W' be a $(1 - .8\frac{1}{2D})$ -viable set for \tilde{T} of size $s = \frac{q}{2D}$, as guaranteed by Lemma 7.2; this set exists provided the condition $s = \Omega(\log q + r \log \frac{q}{s})$ is satisfied. Setting $S = \cup_{i=1}^D A_i W'$ for $\{A_i\}$ as in Lemma 7.3, the resulting set S is of size $\frac{q}{2}$ and by Lemma 6.2 is $(\frac{2D}{8})^2 \Delta$ -viable for \tilde{T} . Using again that \tilde{T} and T are mutually .005-close, Lemma 3.7 implies that S is $2((\frac{2D}{8})^2 \Delta + .005)$ -viable for T . Finally, the condition $D^2 \Delta < \frac{1}{1400}$ from Lemma 7.3 yields that S is .015-viable for T . Since the size of S is $q/2$ we have derived a contradiction with the assumption of minimality of q , under the condition that

$$s = \frac{q}{2D} = \Omega(\log q + r \log \frac{q}{s}) = \Omega(\log q + r \log D),$$

which will hold provided $q = \tilde{\Omega}(r \log D) = \Omega(re^{\tilde{O}(\frac{1}{\gamma} \log^3 d)})$. \square

With Theorem 7.1 in hand, we direct the reader to Appendix B, and specifically the proof of Corollary B.1. Picking up the proof of that corollary at the third sentence with the replacement of $W = V_1^{\log n-1} \otimes V_2^{\log n-1}$ with $W = S$ brings us in line with the remaining results of that section, namely Corollaries B.1, B.2 and B.3, yielding the same results as stated with a quadratic improvement on the dependence on r , from r^2 to r .

Acknowledgements. We thank Andras Molnar for comments on an earlier draft of this paper, and Christopher T. Chubb for comments and the permission to include the suggestive pictures representing the tensor network structure of the isometry produced by our algorithms.

I. Arad's research was partially performed at the Centre for Quantum Technologies, funded by the Singapore Ministry of Education and the National Research Foundation, also through the Tier 3 Grant random numbers from quantum processes. Z. Landau and U. Vazirani acknowledge support by ARO Grant W911NF-12-1-0541, NSF Grant CCF-1410022 and Templeton Foundation Grant 52536 T. Vidick was partially supported by the IQIM, an NSF Physics Frontiers Center (NFS Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028).

References

- [AAI10] Dorit Aharonov, Itai Arad, and Sandy Irani. Efficient algorithm for approximating one-dimensional ground states. *Phys. Rev. A*, 82:012315, Jul 2010.
- [AAVL11] Dorit Aharonov, Itai Arad, Umesh Vazirani, and Zeph Landau. The detectability lemma and its applications to quantum hamiltonian complexity. *New Journal of Physics*, 13(11):113043, 2011.

- [AKL14] Itai Arad, Tomotaka Kuwahara, and Zeph Landau. Connecting global and local energy distributions in quantum spin models on a lattice. Technical report, arXiv:1406.3898, 2014.
- [AKLV13] Itai Arad, Alexei Kitaev, Zeph Landau, and Umesh Vazirani. An area law and sub-exponential algorithm for 1D systems. In *Proceedings of the 4th Innovations in Theoretical Computer Science (ITCS)*, 2013.
- [ALV12] Itai Arad, Zeph Landau, and Umesh Vazirani. Improved one-dimensional area law for frustration-free systems. *Phys. Rev. B*, 85:195145, May 2012.
- [BHM10] Sergey Bravyi, Matthew B. Hastings, and Spyridon Michalakis. Topological quantum order: stability under local perturbations. *Journal of mathematical physics*, 51(9):093512, 2010.
- [CF15] Christopher T. Chubb and Steven T. Flammia. Computing the Degenerate Ground Space of Gapped Spin Chains in Polynomial Time. *ArXiv e-prints*, February 2015.
- [dBOE10] Niel de Beaudrap, Tobias J Osborne, and Jens Eisert. Ground states of unfrustrated spin hamiltonians satisfy an area law. *New Journal of Physics*, 12(9):095007, 2010.
- [DG03] Sanjoy Dasgupta and Anupam Gupta. An elementary proof of a theorem of Johnson and Lindenstrauss. *Random structures and algorithms*, 22(1):60–65, 2003.
- [ECP10] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82(1):277–306, Feb 2010.
- [Fey82] Richard P. Feynman. Simulating physics with computers. *International journal of theoretical physics*, 21(6):467–488, 1982.
- [Has04] Matthew B. Hastings. Lieb-schultz-mattis in higher dimensions. *Physical Review B*, 69(104431), 2004.
- [Has06] Matthew B. Hastings. Solving gapped hamiltonians locally. *Physical review b*, 73(8):085115, 2006.
- [Has07] Matthew B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.
- [Has10] Matthew B. Hastings. Locality in Quantum Systems. *ArXiv e-prints*, August 2010.
- [Hua14] Yichen Huang. A polynomial-time algorithm for the ground state of one-dimensional gapped Hamiltonians. *ArXiv e-prints*, June 2014.
- [Hua15] Yichen Huang. A simple efficient algorithm in frustration-free one-dimensional gapped systems. *ArXiv e-prints*, October 2015.
- [KAAV15] T. Kuwahara, I. Arad, L. Amico, and V. Vedral. Local reversibility and entanglement structure of many-body ground states. *ArXiv e-prints*, arXiv:1502.05330v2, February 2015.
- [KGK⁺14] M. Kliesch, C. Gogolin, J. Kastoryano, M. A. Riera, and J. Eisert. Locality of temperature. *Phys. Rev. X*, 4:031019, Jul 2014.

- [KLW15] JP Keating, N Linden, and HJ Wells. Spectra and eigenstates of spin chain hamiltonians. *Communications in Mathematical Physics*, 338(1):81–102, 2015.
- [LVV15] Zeph Landau, Umesh Vazirani, and Thomas Vidick. A polynomial time algorithm for the ground state of one-dimensional gapped local hamiltonians. *Nature Physics*, 2015.
- [Mas09] Lluís Masanes. Area law for the entropy of low-energy states. *Physical Review A*, 80(5):052104, 2009.
- [MSVC15] Andras Molnar, Norbert Schuch, Frank Verstraete, and J. Ignacio Cirac. Approximating gibbs states of local hamiltonians efficiently with projected entangled pair states. *Phys. Rev. B*, 91:045138, Jan 2015.
- [ÖR95] Stellan Östlund and Stefan Rommer. Thermodynamic limit of density matrix renormalization. *Physical review letters*, 75(19):3537, 1995.
- [ref15] Quantum Hamiltonian Complexity Reunion Workshop, Simons Inst. for the Theory of Computing, UC Berkeley. <https://simons.berkeley.edu/workshops/schedule/3063>, May 2015.
- [RÖ97] Stefan Rommer and Stellan Östlund. Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Physical Review B*, 55(4):2164, 1997.
- [SC10] Norbert Schuch and J. Ignacio Cirac. Matrix product state and mean-field solutions for one-dimensional systems can be found efficiently. *Phys. Rev. A*, 82:012314, Jul 2010.
- [Ver10] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*, 2010.
- [Vid08] G. Vidal. Class of quantum many-body states that can be efficiently simulated. *Phys. Rev. Lett.*, 101:110501, Sep 2008.
- [Vid09] Guifre Vidal. Entanglement renormalization: an introduction. *arXiv preprint arXiv:0912.1651*, 2009.
- [Whi92] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.
- [Whi93] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, Oct 1993.
- [Wil75] Kenneth G Wilson. The renormalization group: Critical phenomena and the kondo problem. *Reviews of Modern Physics*, 47(4):773, 1975.

A Missing proofs

A.1 Viable sets and lemmas on close subspaces

In this section we give proofs for the lemmas from Section 3.1. The proof of Lemma 3.5 follows directly from the following general operator facts:

Lemma A.1. 1. If X and Y are positive operators and $X \geq Y$ then $\text{range}(Y) \subset \text{range}(X)$.

2. If $PQP \geq cP$ for projections P, Q then for every $v \in \text{range}(P)$ of unit norm, there exists $w \in \text{range}(Q)$, $\|w\| = 1$ such that $Pw = c_v v$ for some constant c_v with $|c_v| \geq c$.

Proof. For 1., suppose $y \in \text{range}(Y)$ and let $y = x + x_\perp$, $x \in \text{range}(X)$, $x_\perp \perp \text{range}(X)$ be the orthogonal decomposition. Since $\langle Xx_\perp, x_\perp \rangle = 0$ it follows that $\langle Yx_\perp, x_\perp \rangle = 0$ and thus $x_\perp \perp \text{range}(Y)$ as well and hence $x_\perp = 0$ and $y = x \in \text{range}(X)$.

For 2., it follows from 1. that if $PQP \geq cP$ then for any $v \in \text{range}(P)$ there exists an $r \in \text{range}(P)$ such that $PQP r = PQR = v$. So then $\langle PQPr, r \rangle \geq c\langle Pr, r \rangle = c\|r\|^2$. But $\langle PQPr, r \rangle = \langle v, r \rangle \leq \|r\|\|v\|$. Putting these two inequalities together along with the assumption that $\|v\| = 1$ yields $\|r\| \leq 1/c$. \square

Proof of Lemma 3.6. Since S_1, S_2 are defined on disjoint sets of qudits, it follows that $P_{S^{(ext)}} = P_{S_1^{(ext)}} P_{S_2^{(ext)}}$, and so

$$P_T P_{S^{(ext)}} P_T = P_T P_{S_1^{(ext)}} P_{S_2^{(ext)}} P_T = P_T P_{S_1^{(ext)}} P_T - P_T P_{S_1^{(ext)}} (\mathbb{1} - P_{S_2^{(ext)}}) P_T.$$

The definition of a viable set implies that $P_T P_{S_1^{(ext)}} P_T \geq (1 - \delta_1) P_T$. In addition,

$$P_T P_{S_1^{(ext)}} (\mathbb{1} - P_{S_2^{(ext)}}) P_T \leq P_T (\mathbb{1} - P_{S_2^{(ext)}}) P_T \leq \delta_2 P_T.$$

Therefore, $P_T P_{S^{(ext)}} P_T \geq (1 - \delta_1 - \delta_2) P_T$. \square

Proof of Lemma 3.7. Notice that $P_A P_B P_A \geq (1 - \delta) P_A$ is equivalent to the statement that $\|P_B |a\rangle\|^2 \geq (1 - \delta)$ for all $|a\rangle \in A$ with $\| |a\rangle \| = 1$. It follows for $|t''\rangle \in T'$ of unit norm, $|t'\rangle = P_{T'} |t''\rangle$ has the property that $\| |t'\rangle \|^2 \geq (1 - \delta')$ and thus $\| |t'\rangle - |t\rangle \| \leq \sqrt{\delta'}$. Similarly $|t\rangle = P_T |t'\rangle$ has the property that $\| |t\rangle \|^2 \geq \| |t'\rangle \|^2 (1 - \delta)$ and thus $\| |t'\rangle - |t\rangle \| \leq \| |t'\rangle \| \sqrt{\delta}$. By the triangle inequality, $\| |t''\rangle - |t\rangle \| \leq \sqrt{\delta'} + \sqrt{\delta}$ and since $|t\rangle \in T$, this implies that the distance between $|t''\rangle$ and T is at most $\sqrt{\nu} + \sqrt{\delta}$, i.e. $\| P_{S_{ext}} |t\rangle \|^2 \geq 1 - (\sqrt{\nu} + \sqrt{\delta})^2 \geq 1 - 2(\nu + \delta)$. As mentioned, this last statement is equivalent to T being a $2(\nu + \delta)$ close to T'' . \square

A.2 Random projection lemmas

Lemma A.2 (Johnson-Lindenstrauss). Let $|u\rangle$ be a vector in a q -dimensional Hilbert space \mathcal{H}_A and suppose $0 < \varepsilon < \frac{1}{2}$. Let V be a random s -dimensional subspace of \mathcal{H}_A . Then

$$\Pr \left(\frac{q}{s} \|P_V |u\rangle\|^2 \geq (1 - \varepsilon) \| |u\rangle \|^2 \right) \geq 1 - e^{-\frac{s\varepsilon^2}{4}}.$$

This result follows from applying Lemma 2.2 of Ref. [DG03] as is done in their proof of Theorem 2.1.

Corollary A.3. Let $|u\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with \mathcal{H}_A a q -dimensional Hilbert space, and suppose $0 < \varepsilon < \frac{1}{2}$. Let V be a random s -dimensional subspace of \mathcal{H}_A . Then

$$\Pr \left(\|P_{V_{ext}} |u\rangle\|^2 \leq (1 - \varepsilon) \frac{s}{q} \| |u\rangle \|^2 \right) \leq q e^{-\frac{s\varepsilon^2}{4}}.$$

Proof. Using the Schmidt decomposition, write $|u\rangle = \sum_{i=1}^{q'} \alpha_i |a_i\rangle \otimes |b_i\rangle$, where we are guaranteed that $q' \leq q$. Then $\|P_{V_{ext}}|u\rangle\|^2 = \sum_{i=1}^{q'} \alpha_i^2 \|P_V|a_i\rangle\|^2$. Applying Lemma A.2 to each $|a_i\rangle$ and using the union bound yields

$$\|P_{V_{ext}}|u\rangle\|^2 \geq (1 - \varepsilon) \sum_i \alpha_i^2 \| |a_i\rangle \|^2 = (1 - \varepsilon) \| |u\rangle \|^2,$$

with probability at least $1 - q' e^{-\frac{s\varepsilon^2}{4}}$. □

A.3 Constructing an MPO for the cluster expansion

In this section we provide the proof of Theorem 6.19 and Theorem 6.20 from Section 6.5.1.

Proof of Theorem 6.19. For an integer $m \geq 1$ we let ρ_m be the summation of $f(w)$ over all words w such that there exists m disjoint intervals, each of length at least r , such that the support of w contains each interval but does not contain the two qudits that lie immediately to the left and right of the interval (we call these two qudits the “boundary” of the interval). Using the inclusion-exclusion principle one can verify that

$$e^{-\beta H} - M_r(H) = - \sum_{m=1}^{\infty} (-1)^m \rho_m. \quad (28)$$

We bound the operator norm of each ρ_m individually. Write $\rho_m = \sum_{I=\{I_1, \dots, I_m\}} \rho_I$, where the summation is over all m -tuples of disjoint intervals I_1, \dots, I_m of length at least r , and each ρ_I contains all those h_w for which the support of w contains each of the intervals I_i but not its boundary and is arbitrary everywhere else. Very roughly, the summation is over at most $n^{2m}/(m!)$ terms. Using that the boundaries are excluded, it is not hard to see that $\rho_I = e^{-\beta H_I} \prod_{j=1}^m \eta(I_j)$, where H_I contains all terms in the Hamiltonian that do not act on the qudits in the boundary of I_j and $\eta(I_j)$ is the sum of all $f(w)$ such that the support of w is exactly I_j . Using $\|e^{-\beta H_I}\| \leq 1$ we can bound

$$\begin{aligned} \|\rho_I\| &\leq \prod_j \|\eta(I_j)\| \\ &\leq \prod_j \left(\sum_{w: \text{supp}(w)=I_j} \frac{(-\beta)^{|w|}}{|w|!} \right) \\ &= (e^\beta - 1)^{\sum_j |I_j|}. \end{aligned}$$

Combining with (28),

$$\begin{aligned} \|e^{-\beta H} - M_r(H)\| &\leq \sum_{m=1}^{\infty} \|\rho_m\| \\ &\leq \sum_{m=1}^{\infty} \sum_{I=\{I_1, \dots, I_m\}} \|\rho_I\| \\ &\leq \sum_{m=1}^{\infty} \frac{n^{2m}}{m!} (e^\beta - 1)^{mr} \\ &= e^{n^2(e^\beta - 1)^r} - 1, \end{aligned}$$

where for the third line we used that β is such that $e^\beta - 1 < 1$. □

Proof of Theorem 6.20. The r^{th} expansion of $e^{-\beta H}$ is given by

$$M_r(H) := \sum_{w \in S_r} \frac{(-\beta)^{|w|}}{|w|!} h_w,$$

where w is a word on the alphabet of local Hamiltonian terms $\{1, \dots, n-1\}$, $h_w := \prod_{i \in w} h_i$, and S_r is the set of words in which all connected components have a support of size at most $r-1$. Let $\underline{I} = (I_1, I_2, \dots, I_m)$ be a collection of disjoint segments on the line, and $\max(\underline{I})$ denote the length of the largest segment in \underline{I} . We write $w \in \underline{I}$ to mean that the connected components of w matches the segments specified by \underline{I} . Using this notation, $M_r(H)$ can be rewritten as

$$M_r(H) = \sum_{\max(\underline{I}) < r} \sum_{w \in \underline{I}} \frac{(-\beta)^{|w|}}{|w|!} h_w.$$

A rather straightforward combinatorial argument shows that for a given $\underline{I} = (I_1, \dots, I_m)$,

$$\sum_{w \in \underline{I}} \frac{(-\beta)^{|w|}}{|w|!} h_w = \prod_{j=1}^m \sum_{w \in I_j} \frac{(-\beta)^{|w|}}{|w|!} h_w,$$

where the notation $w \in I_j$ means that the support of the word w has a single connected component whose support is I_j . Therefore, if we define for each segment I

$$\rho_I := \sum_{w \in I} \frac{(-\beta)^{|w|}}{|w|!} h_w, \quad (29)$$

then

$$M_r(H) = \sum_{\max(\underline{I}) < r} \rho_{I_1} \otimes \rho_{I_2} \otimes \dots \otimes \rho_{I_m}. \quad (30)$$

We use Eq. (30) as the basis for an efficient MPO representation of $M_r(H)$.

1st step: creating a table of ρ_I . The first step is a pre-processing step, which can be run performed before the start of the algorithm. Its goal is to create a table of MPO representations of *all* ρ_I that appear in Eq. (30). This can be done in $nd^{O(r)}$ time. Indeed, note first that the total number of intervals I to consider is at most nr . The associated MPO can be computed iteratively, starting with $I = \emptyset$ for which $\rho_\emptyset = \mathbb{1}$. Assuming all ρ_I with $|I| < s$ have been determined, compute an MPO for ρ_I , for any I such that $|I| = s$, as follows. Clearly,

$$\rho_I = e^{-\beta H_I} - \sum_{\underline{I}'} \rho_{I'_1} \otimes \rho_{I'_2} \otimes \dots \otimes \rho_{I'_m},$$

where the summation runs over all disjoint subsets $\underline{I}' = (I'_1, I'_2, \dots, I'_m)$ included in I and with $m \geq 2$. An MPO for the first term can be obtained in time $d^{O(s)}$ by direct matrix exponentiation. The second term is expressed as the sum of most 2^s terms, for each of which an MPO was computed in a previous iteration. Altogether ρ_I can therefore be computed in time $d^{O(s)}$ and stored in memory as an MPO of bond dimension at most d^r .

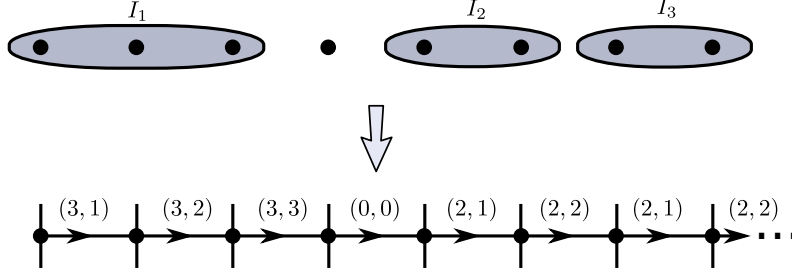


Figure 2: An example of the (ℓ, k) indices that give rise to the configuration of segments $I_1 = (1, 2, 3)$; $I_2 = (5, 6)$; $I_3 = (7, 8)$.

2nd step: creating the MPO of $M_r(H)$ We follow the expansion Eq. (30), using a signaling mechanism through which every site tells the site to its right to which ρ_I it belongs. This ensures that every non-vanishing contraction of the virtual indices corresponds to exactly one product $\rho_{I_1} \otimes \rho_{I_2} \otimes \dots$ from Eq. (30).

Virtual bonds are indexed by triples (ℓ, k, α) . The virtual bond across sites $a, a+1$ describes the segment I to which a belongs: $\ell \in \{0, 1, \dots, r-1\}$ denotes the width of I , $k \in \{1, \dots, r-1\}$ denotes the position of the site a within I , and α corresponds to the index of the virtual bond in the MPO expansion of ρ_I . For example, suppose that site a is in third position in the support of ρ_I , where $|I| = 8$. Then it transmits to site $a+1$ the indices $\ell = 8, k = 3$. Site $a+1$ will then transmit to $a+2$ the indices $\ell = 8, k = 4$ and so on. When the last site in ρ_I is reached, in our example site $a+5$, it transmits to $a+6$ the indices $(k = 8, \ell = 8)$. Then $a+6$ could either be an empty site, transmitting $\ell = k = 0$ to the right, or start a new segment I with any $\ell > 0$. See Fig. 2 for an illustration of this signaling mechanism.

To write a formal definition of the MPO, let us use $[A^{(a)}(I)]_{\alpha_1, \alpha_2}^{i, j}$ to denote the tensor associated with ρ_I at site $a \in I$. In order to simplify notation, when the site a is the left-most (resp. right-most) site in I we use the convention that $[A^{(a)}(I)]_{\alpha_1, \alpha_2}^{i, j}$ is non-vanishing only when $\alpha_1 = 1$ (resp. $\alpha_2 = 1$). Finally, we denote each segment I by $I(\ell, a)$ where ℓ is the width of the segment and a is its first site. For a non-extremal site a , the tensor $A^{(a)}$ of $M_r(H)$ is given by

$$[A^{(a)}]_{(\ell_1, k_1, \alpha_1), (\ell_2, k_2, \alpha_2)}^{i, j} := \begin{cases} [A^{(a)}(I(\ell_1, a - k_1 + 1))]_{\alpha_1, \alpha_2}^{i, j} & \text{for } k_1 < \ell_1 \text{ and } \ell_1 = \ell_2, \text{ and } k_2 = k_1 + 1, \\ [A^{(a)}(I(\ell_2, a))]_{\alpha_1, \alpha_2}^{i, j} & \text{for } k_1 = \ell_1 \text{ and } 0 < \ell_2 \leq n - a + 1 \text{ and } k_2 = 1, \\ \delta_{i, j} & \text{for } k_1 = \ell_1 \text{ and } \ell_2 = k_2 = 0 \text{ and } \alpha_1 = \alpha_2 = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (31)$$

The first case corresponds to a site a in the interior of the segment $I = I(\ell_1, a - k_1 + 1)$. The second case corresponds to an a that is the first site of a new segment $I = I(\ell_2, a)$. Note that the condition $\ell_2 \leq n - a + 1$ guarantees that this segment does not exceed the right side of the chain. Finally, the third case corresponds to an empty site a .

To complete the definition it remains to specify $A^{(1)}$ and $A^{(n)}$. Just as the tensors for ρ_I , we keep both left and right indices but make them non-zero only when $\ell = k = 0$ and $\alpha = 1$. Then $A^{(1)}$ is defined as $A^{(a)}$ with the additional requirement that it is non-vanishing only when $\ell_1 = k_1 = 0$ and $\alpha_1 = 1$. The tensor $A^{(n)}$ is defined directly by (31). In that case, for every (ℓ_1, p_1, α_1) there is at most one triple (ℓ_2, p_2, α_2) for which $A^{(n)}$ is non-vanishing, and so without loss of generality we can map it to $\ell_2 = k_2 = 0$ and $\alpha_2 = 1$.

To finish the proof note that the virtual bond dimension is bounded by $r(r-1)d^r < r^2 d^r$, and therefore the second step can be done in time $nd^{O(r)}$ since it only involves local assignments. \square

B Proof of bounds for the area law

In this section we prove the three main claims made in Corollary 3.10. For convenience we restate the three claims as three separate corollaries.

Corollary B.1 (Schmidt Rank of Ground Space). *Let H be a local Hamiltonian satisfying Assumption (DG). For any cut and any $\delta = \text{poly}^{-1}(n)$ there is subspace $S \subseteq \mathcal{H}$ that is δ -close to T and such that every element of S has Schmidt Rank no larger than $s(\delta) = \tilde{O}(r^2)e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)} \cdot e^{\tilde{O}(\gamma^{-1/4} \log^{3/4}(\frac{1}{\delta}) \log d)}$.*

Proof. For simplicity assume the cut we are interested in is the middle cut. Recall that the two sets $V_1^{\log n-1}$ and $V_2^{\log n-1}$ coming from the second to last iteration in Procedure 1 each have dimension $q = \tilde{O}(r^2)e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}$ and are .015-viable for T . Thus by Lemma 3.6 the set $W = V_1^{\log n-1} \otimes V_2^{\log n-1}$ is .03-viable for T . The tensor product structure ensures that every element of W has Schmidt rank no larger than q . Invoking Corollary 6.14 and proceeding as in the proof of Theorem 3.8 but with $\ell = \Theta(t^{3/4})$ where $t = \Theta(\log \delta^{-1})$ we obtain a spectral AGSP K with

$$D = e^{\tilde{O}(\gamma^{-1/4} \log^{3/4}(\frac{1}{\delta}) \log d)}, \quad \Delta \leq \delta/2$$

for a Hamiltonian \tilde{T} such that $\tilde{T} = \tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \frac{1}{10}\gamma]}$ and T are mutually $(\delta/2)$ -close. From Lemma 6.2 the set $W_\delta = KW$ is $(\delta/2)$ -viable for \tilde{T} and every element within it has Schmidt rank no larger than qD . Since \tilde{T} and T are $(\delta/2)$ -close, W_δ is δ -viable for T . \square

Corollary B.2 (Area Law for Ground Space). *Let H be a local Hamiltonian satisfying Assumption (DG). Then across every cut, every element $|\Gamma\rangle$ of the ground space of H has entanglement entropy*

$$S(|\Gamma\rangle\langle\Gamma|) \leq 4 \log r + \tilde{O}\left(\frac{1}{\gamma} \log^3 d\right).$$

Proof. In light of the values of $s(\delta)$ in Corollary B.1, the proof is standard and follows the line of Ref. [AKLV13]. We bound

$$S(|\Gamma\rangle\langle\Gamma|) \leq \log\left(\tilde{O}(r^2)e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)}\right) + \sum_{i=3}^{\infty} 2^{-i} \log(s(2^{-(i+1)})),$$

which is dominated by the first term. \square

Corollary B.3 (MPS representation for Ground Space). *Let H be a local Hamiltonian satisfying Assumption (DG). For any $\delta = \text{poly}^{-1}(n)$ and any state $|\Gamma\rangle$ in the ground space of H , there exists a state $|\Gamma'\rangle$ with $|\langle\Gamma|\Gamma'\rangle| > 1 - \delta$ such that $|\Gamma'\rangle$ has an MPS representation with bond dimension at most B , where*

$$B = \tilde{O}(r^2)e^{\tilde{O}(\frac{1}{\gamma} \log^3 d)} e^{\tilde{O}(\gamma^{-1/4} \log^{3/4}(\frac{n}{\delta}) \log d)}.$$

Proof. To show the existence of a matrix product state of bond dimension B within δ of $|\Gamma\rangle$ it suffices to show the existence for every cut of a state of entanglement rank B within $\frac{\delta}{n}$ of $|\Gamma\rangle$, as it is not hard to see that the errors made at each cut will add up linearly. Thus Corollary B.1 yields the existence of $|\Gamma'\rangle$ with an MPS representation with bond dimension bounded by $s(\delta/n) = e^{\tilde{O}(\gamma^{-1/4} \log^{3/4}(\frac{n}{\delta}) \log d)}$. \square

We sketch how the dependence on r can be improved from r^2 to r in all three bounds above. First, the bound on the Schmidt rank from Corollary B.1 can be improved, replacing r^2 with $r^{1+\epsilon}$, by using AGSPs with a D, Δ -tradeoff of the form $D^{1/\epsilon} \Delta \leq \frac{1}{2000}$, instead of $D^{12} \Delta \leq \frac{1}{2000}$. This can be further improved to $r^{1+o(1)}$ by directly substituting the bound Δ desired for error reduction and deducing the correct value of D from the AGSP construction, rather than fixing a D, Δ -tradeoff and using a suitable power of the AGSP to achieve the desired error reduction. We do not provide more details as the shorter “bootstrapping” argument given in Section 7 yields a similar improvement.

C Trimming

In this section we describe the trimming mechanism that underlies the procedure **Merge**’ used in Algorithm 1.

The trimming procedure introduced in previous work [LVV15] relied on the observation that given a good approximation to a target vector $|v\rangle$ of low bond dimension, trimming the approximating vector by dropping Schmidt vectors associated with the smallest Schmidt coefficients at each cut yields an almost-as-good approximation to $|v\rangle$ with lower bond dimension. In the present scenario the approximating vector is not known: instead we are given a basis for a subspace that contains the approximating vector. The natural idea consists in trimming MPS representations for the basis vectors in a way that guarantees that $|v\rangle$ is still closely approximated by some vector in the span of the resulting set. It is not clear if independently trimming each of the basis vectors, as done in Ref. [LVV15] works – indeed, the basis vectors themselves could a priori have a very flat “Schmidt spectrum” (meaning trimming would induce large changes), while $|v\rangle$ still does not.

We provide a slightly modified procedure which starts with a basis for the viable set and trims the basis vectors collectively at each cut, from the rightmost to the leftmost cut, as follows: for each cut, project each element of the basis onto the span of the left Schmidt vectors of *any* basis element that is associated with a large Schmidt coefficient.

Definition C.1 (Trimming). *Let $S \subseteq \mathcal{H}_1$ be a δ -viable set for $T \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$ specified by an orthonormal basis $\{|u_i\rangle, i = 1, \dots, s\}$. Suppose $\mathcal{H}_1 = \mathcal{H}_1^1 \otimes \dots \otimes \mathcal{H}_1^k$ for some k . Let $|\psi\rangle = \sum_i |u_i\rangle |i\rangle$. For j from $(k-1)$ down to 1 define $P_{\geq \xi}^j$ inductively as the projection on the span of the left Schmidt vectors of $P_{\geq \xi}^{j+1} \dots P_{\geq \xi}^{k-1} |\psi\rangle$ across the $(j : j+1)$ cut with associated Schmidt coefficient at least ξ . Then the ξ -trimmed set is*

$$\text{Trim}_\xi(S) := \text{Span} \left\{ \left((P_{\geq \xi}^1 \otimes \mathbb{1}_{\mathcal{H}_1^2 \otimes \dots \otimes \mathcal{H}_1^k}) \dots (P_{\geq \xi}^{k-1} \otimes \mathbb{1}_{\mathcal{H}_1^k}) (P_{\geq \xi}^k) \right) |u_i\rangle, i = 1, \dots, s \right\}. \quad (32)$$

With this notion of trimming, we show that if a set S is a good viable set for a set T whose elements are guaranteed to have low bond dimension (e.g. they satisfy an area law) then the result of trimming the set S does not degrade the quality of the viable set too much.

Lemma C.2 (Trimming). *Let $S \subseteq \mathcal{H}_1$ be a δ -viable set for $T \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$ of dimension s . Suppose $\mathcal{H}_1 = \mathcal{H}_1^1 \otimes \dots \otimes \mathcal{H}_1^\ell$ for some ℓ . Let r be an upper bound on the Schmidt rank of any vector in T across any cut $(j : j+1)$ for $j = 1, \dots, \ell-1$. Then the ξ -trimmed set $\text{Trim}_\xi(S)$ is a δ' -viable set for T for $\delta' \leq \delta + 2\sqrt{\ell r s \xi}$.*

Furthermore, a spanning set for $\text{Trim}_\xi(S)$ containing at most s vectors of Schmidt rank at most $s\xi^{-2}$ across any cut can be computed in time $O(\ell M(dsq))$, where q is an upper bound on the dimension of MPS representations for a basis of S and $M(\cdot)$ denotes matrix multiplication time.

Proof. Let $|v\rangle \in T$ and $|u\rangle = \sum_i \mu_i |u_i\rangle |v_i\rangle$ a unit vector such that $|\langle u|v\rangle|^2 \geq 1 - \delta$. For $j = 0, \dots, \ell$ let

$$|u'_j\rangle = (P_{\geq \xi}^j \otimes \mathbb{1}_{\mathcal{H}_1^{j+1} \otimes \dots \otimes \mathcal{H}_1^\ell} \otimes \mathbb{1}_{\mathcal{H}_2}) \cdots (P_{\geq \xi}^\ell \otimes \mathbb{1}_{\mathcal{H}_2}) |u\rangle,$$

and for $i \in \{1, \dots, s\}$,

$$|u_i^j\rangle = (P_{\geq \xi}^j \otimes \mathbb{1}_{\mathcal{H}_1^{j+1} \otimes \dots \otimes \mathcal{H}_1^\ell}) \cdots (P_{\geq \xi}^\ell) |u_i\rangle.$$

By definition the Schmidt coefficients of the vector $((\mathbb{1} - P_{\geq \xi}^j) P_{\geq \xi}^{j+1} \cdots P_{\geq \xi}^{\ell-1} \otimes \mathbb{1}) |\psi\rangle$, where $|\psi\rangle = \sum |u_i\rangle |i\rangle$, across the cut $(j, j+1)$ are all at most ξ . Since acting with a local projection (here, $|i\rangle\langle i|$ on \mathcal{H}_2) cannot increase the largest Schmidt coefficient, the same holds of the vector $((\mathbb{1} - P_{\geq \xi}^j) \otimes \mathbb{1}) |u_i^{j+1}\rangle$. Based on these observations we may upper bound, for any i, ℓ ,

$$|\langle u_i^{j+1} | \langle v_i | ((\mathbb{1} - P_{\geq \xi}^j) \otimes \mathbb{1}_{\mathcal{H}_1^{j+1} \otimes \dots \otimes \mathcal{H}_1^\ell} \otimes \mathbb{1}_{\mathcal{H}_2}) | a_\ell \rangle | b_\ell \rangle| \leq \xi,$$

where the inequality follows since we are taking the inner product of a vector with largest Schmidt coefficient at most ξ with another vector of Schmidt rank 1. Using the promised bound on the Schmidt rank of $|v\rangle$ we deduce

$$|(\langle u'_j | - \langle u'_{j+1} |) v\rangle| \leq \xi \sqrt{rs}.$$

Using that the $(|u'_j\rangle - |u'_{j+1}\rangle)$ are orthogonal for different values of j (which follows since we may assume without loss of generality that the projections $P_{\geq \xi}^j$ commute), we deduce that the projection of $|v\rangle$ on $\text{Span}\{|u'_j\rangle - |u'_{j+1}\rangle, j = 1, \dots, \ell-1\}$ has squared norm at most $\ell \xi^2 rs$. Thus in particular $|\langle v | (u'_1 - u'_\ell) \rangle| \leq \xi \sqrt{2\ell rs}$ and the claimed bound on δ' follows.

For the “furthermore” part, note that $|\psi\rangle$ has at most s/ξ^2 Schmidt coefficients larger than ξ across any cut $(j : j+1)$. The application of each $P_{\geq \xi}^j$ reduces the Schmidt rank across the cut $(j : j+1)$ to at most s/ξ^2 , while never increasing it at any of the previously considered cuts. The left Schmidt vectors of

$$(P_{\geq \xi}^1 \otimes \mathbb{1}_{\mathcal{H}_1^2 \otimes \dots \otimes \mathcal{H}_1^\ell}) \cdots (P_{\geq \xi}^{\ell-1} \otimes \mathbb{1}_{\mathcal{H}_1^\ell}) (P_{\geq \xi}^\ell) |\psi\rangle$$

across the cut specified by the division $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ form a spanning set for $\text{Trim}_\xi(S)$.

In order to compute canonical MPS representations for a basis of $\text{Trim}_\xi(S)$ we first create an MPS representation for $|\psi\rangle$ and reduce it to canonical form. This costs $O(\ell M(dsq))$ operations, where $M(\cdot)$ is matrix multiplication time, and $M(dsq)$ is the time required to perform required basic operations on tensors of bond dimension $O(dsq)$, such as singular value decompositions. Proceeding from the cut $(\ell-1, \ell)$ to the $(1, 2)$ cut from right to left, we then set the coefficients of the diagonal tensor matrices Λ_j from the MPS representation that are smaller than ξ to zero. The resulting re-normalized state is automatically given in canonical MPS form, and a spanning set for $\text{Trim}_\xi(S)$ can be obtained by cutting the last bond. \square